Universidad Autónoma de Madrid
Departamento de Biología Molecular
Facultad de Ciencias

Structure-based characterization of the multi-modular Developmentally Regulated GTPase in complex with the DRG family regulatory protein

Memoria presentada por Sandrea Maureen Francis

para optar al grado de Doctora en Ciencias Biológicas

Madrid, Octubre de 2012

El trabajo presentado en esta memoria fue realizado en el Instituto de Biomedicina de Valencia (IBV-CSIC), bajo la dirección del Dr. Jerónimo Bravo Sicilia.

Dr. Jerónimo Bravo Sicilia,

CERTIFICA

Que la tesis doctoral titulada “Structure-based characterization of the multi-modular Developmentally Regulated GTPase in complex with the DRG family regulatory protein”, ha sido realizada por Dª. Sandrea Maureen Francis, bajo su dirección en el Instituto de Biomedicina de Valencia (IBV-CSIC) y que, salvo mejor criterio del tribunal que ha de juzgarla, reúne todos los requisitos de originalidad y calidad necesarios para optar al grado de Doctora en Ciencias Biológicas.

Y para que conste, expido el presente certificado en Valencia a 10 de Septiembre de 2012

Fdo. Dr. Jerónimo Bravo Sicilia

A mí família...

Acknowledgements

It is a pleasure to thank the many people who have made this thesis possible with their help and support.

I would like to express my heartfelt gratitude to my thesis supervisor Dr. Jerónimo Bravo for giving me this opportunity to develop this doctoral thesis. I deeply appreciate the countless educational, insightful, motivating and inspirational discussions that we used to have. I am thankful to him for pointing me in the right direction in the course of this work while still giving the freedom to think independently and for his support throughout the thesis, approachability, kindness and fun time.

I have been blessed with a totally loveable and cheerful lot in my daily work who have made coming to the lab every morning joyful and made every day an enjoyable experience. I would like to thank Isa, Leti, Marcin and Sara, my colleagues, for all their constant love, support and help through difficult and good times during the development of this work.

I express my deepest gratitude to Dr. Bertrand Séraphin for the collaborative work and for allowing me to do a short term stay at the IGBMC. I am profoundly indebted to Maria for taking the time and patience to teach me yeast experimental biology in a short time and for helping me have a wonderful stay in Strasbourg. I also thank all the lab members at the IGBMC for their help and sharing of good times.

I wish to thank my former colleagues of CNIO, Jose, Lina, Nayra and Rebecca; Mer for just being the kind, generous and helpful girl that she is. I owe her so much that can never be repaid. I am grateful to Miguel for writing me the recommendation letter which indeed turned out to be very rewarding. I sincerely thank Miguel and Katrin (NIMR, MRC) for kindly taking the time to evaluate this thesis work. It is difficult to overstate my gratitude to Fede who has helped me take the baby steps to protein crystallography. I am indebted to him for teaching me so much and for being a delight working with. I thank Carme who was instrumental in my arrival to Spain and I am honoured to have had the opportunity to know a great scientist like Ángel. His perseverance and strength till the end is deeply admired. I thank all the guys of the CBM, Antonio for the guidance and help, David for being so cool with computers, Alberto, Raul, Ugo, Javi and the rest of the lab members. I thank Ruben for just being “flipado” and hilarious and for being a good friend who can put a smile on your face anytime of the day. I will remember their friendships and cherish those memories forever.

I would like to take this opportunity to acknowledge everyone at my institute. I thank Aaron, Alberto, Laura, Marisa and Patricia; Paqui for the innumerable crystallization plates and also Clara and Jordi for doing an important step in a structural biologists work. I am most grateful to Marta, Ana and Sole for welcoming me in their lab during the initial months of my arrival at the IBV. It was particularly kind of everyone at the laboratory of Pascual Sanz for helping me and I am profoundly indebted

to Ada for all the support with yeast work. I acknowledge all the other members at the institute and the guys at the administration, library, sanitation, security and maintenance facilities who painfully ensured the proper running of the institute and made it an enjoyable place to work in.

I am extremely grateful to the laboratory of Maria Eugenia, especially Silvia for kindly teaching me the activity assays for GTPases. I offer my sincerest gratitude to the staff at the ESRF, the local contacts and scientists, Sylvia, Hassan and Richard Kahn. I especially thank Sylvia who helped with the crystal dehydration experiments and for being so helpful and welcoming.

I would like to take this opportunity to remember all my friends from school, college, university and NIPER for their love, camaraderie and motivation. I thank my former flat mate Eva, for her friendship and care, and for giving me those Christmas family dinners away from home. I am grateful to Angie and Irene for being amazing “compis de piso”. I would also like to thank the indian community at large in Madrid for providing those special indian moments, specially Rashmi, Amit and Jalaj for all the good and special times that I had with them.

I deeply thank God Almighty for carrying me forward through the paths strewn with stones and roses. I dedicate this thesis to my parents, who have always being there for me, instilling in me the motivation for creative pursuits and without whom I would not be, what I am or where I am. I am indebted to my siblings, whose love and affection has always been a strong pillar of support.

I also wish to thank the countless anonymous people around me who have helped me in some sort without either of us realizing. I am deeply sorry but extremely thankful to all the persons who might have been left out due to limitations of memory.

Index

Figure and table index

Figures

Tables

Summary

Summary

Developmentally **R**egulated **G**TP binding protein (DRG) subfamily is an uncharacterized member of the OBG family of the TRAnslation FACtor (TRAFAC) class of GTPases, highly conserved evolutionarily in the two major superkingdoms, archaea and eukaryota. Absent in eubacteria, two related proteins, Drg1 and Drg2, are present in all eukaryotic organisms where they accumulate only in the presence of DRG Family Regulatory Protein (DFRP): with Dfrp1 binding specifically to Drg1 and Dfrp2 preferentially binding to Drg2. They are thought to play, along with their binding partners, an important role in cell growth and differentiation. A direct involvement in translation was earlier observed when the yeast Drg1 homologue, named as Rbg1 for **R**ibosome **B**inding **G**TPase, in complex with Tma46, the **T**ranslation **M**achinery **A**ssociated protein 46 (yeast Dfrp1 homologue) associated with translating ribosomes.

In the present study, the X-ray crystallographic structure of the yeast Rbg1 GTPase in complex with the C-terminal region of its DFRP partner, Tma46, is reported for the first time. The Rbg1-Tma46 complex structure reveals that DRG proteins are multi-modular factors containing unreported domains similar to those found evolutionarily conserved in proteins directly related to translation. Rbg1 was found to contain the well-known GTP binding domain typical to GTPases and the functionally unknown C-terminal TGS domain. Interestingly, two hitherto-unknown domains were also discovered, one of them a helix-turn-helix and the other displaying an unusual βαβ fold with similarity to the C-terminal domain of the Ribosomal S5 protein. We have named these novel domains of Rbg1 as the HTH and S5D2L domains respectively. They pack adjacent to each other forming a globular unit protruding from the core GTP binding domain. Surprisingly in fact, the S5D2L domain is inserted in the middle of the G-domain between the G3 and G4 conserved GTP binding sequences. Contrary to previous reports, we found that the interaction surface of Tma46 with Rbg1 is far more extensive. The Tma46 C-terminal dfrp fragment was seen to be highly unstructured comprising helical structures and wrapping itself around the G-domain and TGS domains of Rbg1.

Analysis of the *in vitro* GTPase activity of the DRGs, also in the context of DFRP proteins demonstrated that the yeast DRG-DFRP and human DRG1-DFRP1 complexes have slow catalytic values similar to the OBG family members, which might be modulated by its various domains as well as interaction with its partner, the DFRP protein. Additionally, the activity was observed in the absence of any GTPase activating proteins (GAP) or Guanine exchange factors (GEF) unlike the case with small monomeric GTPases. This phenomenon was also suggested earlier. RNA homopolymer assays revealed the non-specific RNA binding capabilities of the DFRPs and the individual domains of DRGs.

Structural information has provided further support to the possible role of the DRG proteins in the translational machinery with the biochemical analyses dissecting the contribution and importance of each of the domains of this multi-modular protein and its interaction with its regulatory protein, DFRP.

Resumen

La subfamilia de las proteínas de unión a GTP reguladas durante el desarrollo (siglas en inglés DRG) es un miembro poco caracterizado de la clase de GTPasas perteneciente a la familia OBG de factores de traducción (TRAFAC), altamente conservadas evolutivamente en dos grandes super-reinos: arqueas y eucariota. Dos proteínas relacionadas, Drg1 y Drg2, ausentes en eubacterias, están presentes en todos los organismos eucariotas donde se acumulan sólo en presencia de la denominada familia de proteínas de regulación de DRG (DFRP): con Dfrp1 uniéndose específicamente a Drg1 y Dfrp2 preferentemente uniéndose a Drg2. Se cree que desempeñan, junto con sus proteínas asociadas, un papel importante en el crecimiento y la diferenciación celular. Se ha observado con anterioridad una participación directa en la traducción, pues el homólogo de Drg1 en levadura, denominado Rbg1 (del inglés "**R**ibosome **B**inding **G**TPase") formando un complejo con Tma46 (del inglés **T**ranslation **M**achinery **A**ssociated protein 46 ", el homólogo en levadura de Dfrp1), localizado con los ribosomas en fase de traducción.

En el presente estudio, se presenta la estructura cristalográfica de rayos X de la GTPasa Rbg1 de levadura en complejo con la región C-terminal de su socio DFRP, Tma46. La estructura del complejo Rbg1-Tma46 revela que las proteínas DRG son factores multimodulares que contienen dominios no descritos previamente, pero similares a otros evolutivamente conservados y que se hallan en proteínas directamente relacionadas con la traducción. Se ha encontrado que Rbg1 contenía el ampliamente descrito dominio de unión a GTP típico de las GTPasas además del dominio C-terminal TGS de función desconocida. Curiosamente, se han identificado además dos dominios desconocidos, uno, una hélice-giro-hélice y el otro, con un plegamiento βαβ inusual con similitud al dominio C-terminal de la proteína ribosomal S5. Estos nuevos dominios de Rbg1 han sido denominados como los dominios HTH y S5D2L respectivamente y empaquetan adyacentes entre sí formando una unidad globular que sobresale desde el dominio de unión al GTP. De hecho, sorprendentemente, el dominio S5D2L se inserta en medio del dominio G entre las secuencias conservadas de unión a GTP G3 y G4. Contrariamente a lo señalado en informes anteriores, se encontró que la superficie de la interacción de Tma46 con Rbg1 es mucho más extensa. El fragmento C-terminal dfrp de Tma46 se encuentra altamente desestructurado compuesto principalmente de estructuras helicoidales que se envuelven alrededor de los dominios G y TGS de Rbg1.

El análisis in vitro de la actividad GTPasa de los DRGs, en el contexto también de las proteínas DFRP ha demostrado que los complejos de levadura DRG-DFRP y humanos DRG1-DFRP1 tienen valores catalíticos bajos similares a otros miembros de la familia de OBG, que pueden ser modulados por sus distintos dominios así como la interacción con su ligando celular, la proteína DFRP. Además, la actividad se observó en la ausencia de proteínas GAP o GEF, como se sugirió con anterioridad, pero de forma distinta a como

sucede con las GTPasas monoméricas pequeñas. Ensayos con homopolímero de ARN revelaron que existe una capacidad no específica de unión del ARN por parte de los DFRPs y algunos de los dominios individuales que forman parte de los DRGs.

La información estructural ha proporcionado un mayor apoyo a la posible función de las proteínas de DRG en la maquinaria de traducción con los análisis bioquímicos, que han diseccionado la contribución y la importancia de cada uno de los dominios de esta proteína multimodular y su interacción con la proteína reguladora, DFRP.

Introduction

1. Introduction

1.1 The GTPase superfamily

GTP binding proteins or G-proteins represent a diverse category of regulatory molecules in the cell and are components of signal transduction pathways in eukaryotes and seemingly also in prokaryotes. G-proteins are characterized by their common ability to bind and hydrolyse guanosine 5′-triphosphate (GTP) to guanosine 5′-diphosphate (GDP) acting as molecular/binary switches between their GTP-bound (active) and GDP-bound (inactive) forms. The GTP binding domain (G-domain) of these proteins contain five characteristic motifs which are called the G1 (Walker A/P-loop), G2 (Switch I), G3 (Walker B/Switch II), G4 and G5. **G1** motif (GxxxxGK(S/T)) is responsible for binding of α- and β-phosphate groups of the nucleotide, **G2** (x(T/S)x) motif is involved in Mg^{2+} binding while **G3** (DxxG) binds to both Mg^{2+} and the γ-phosphate, **G4** ((N/T)KxD) binds directly with the nucleotide through the residues K and D and the weakly conserved G5 is involved basically in guanine base recognition. The P-loop GTPases, called so by virtue of the signature Walker A motif between a helix and a strand (Walker *et al.,* 1982), function by producing conformational changes between the GTP/GDP bound forms of the G-protein transducing cellular signals to downstream effector proteins mainly through changes in switch I (G2) & II (G3) regions (Bourne *et al.,* 1990, Bourne *et al.,* 1991, Leipe *et al.,* 2002, Wittinghofer and Vetter 2011).

Enzymes that used GTP as opposed to ATP were first characterised as initiation, elongation and release factors in the process of protein synthesis in a cell; Elongation factor-Tu (EF-Tu) being the first three dimensional structure to be obtained of a GTPase (Kabsch *et al.,* 1977). This was followed by the identification of the oncogenic Ras family of proteins involved in several processes like gene expression (Ras, Rho), cytoskeleton reorganization (Rho, Rac), nuclear import (Ran), vesicle transport (Rab, Sar1/Arf) and mitogenesis (Ras, Rap) (Garcia-Ranea and Valencia 1998, Macara *et al.,* 1996). Structural and genomic-sequence comparison studies have identified numerous other GTPases having roles in diverse biological/cellular processes, the active form regulating a broad palette of cellular events. Accordingly, the GTPase superclass has been divided into two classes, the TRAFAC (for TRAnslation FACtor-related) class which contains majority of the GTPases and the SIMIBI (for Signal recognition GTPases, MinD and BioD superfamily) class (Leipe *et al.,* 2002, Verstraeten *et al.,* 2011). Further divisions of the TRAFAC class are represented in figure 1.1.

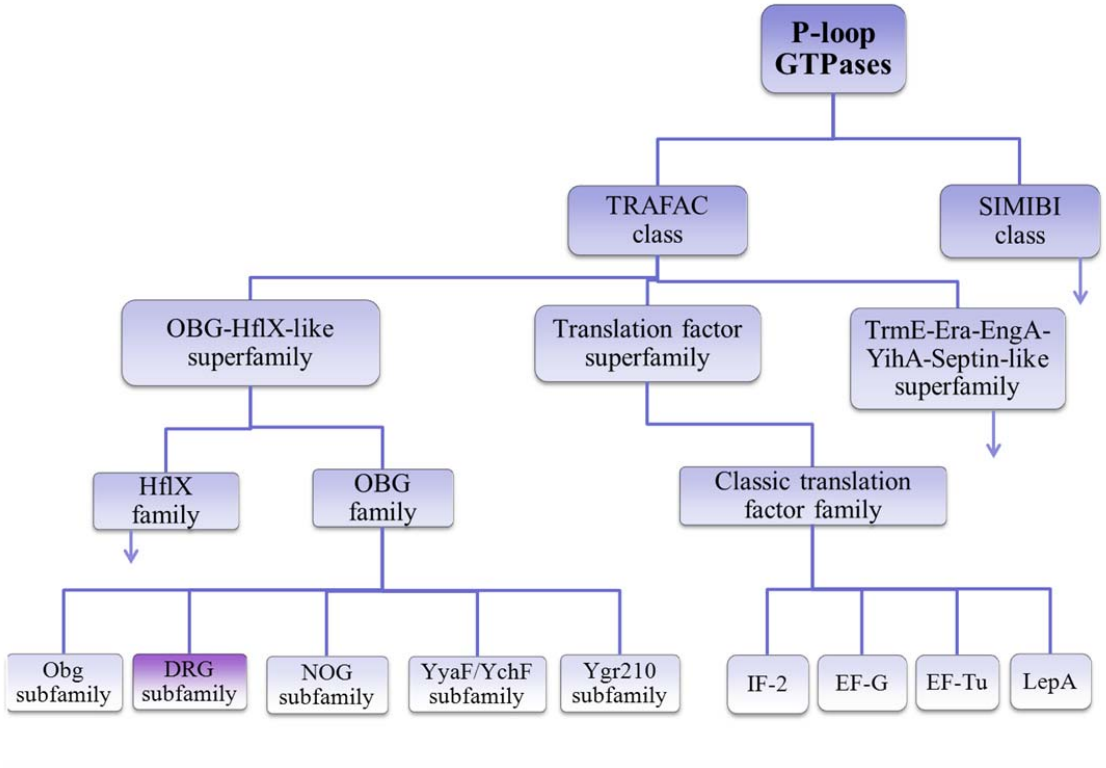


Figure 1.1: Classification of the P-loop GTPases. The family classification tree of the TRAFAC class of GTPases is depicted focusing on OBG family with some of the branches not relevant in this study being not included. The figure is modified from Verstraeten *et al.,* 2011.

The classic translation factor family is spread in all three superkingdoms (Archaea, Bacteria and Eukaryota) and might be vertically inherited from the Last Universal Common Ancestor (LUCA) of all life forms. They include GTPases involved in the translation machinery. The initiation factor GTPases are involved in the formation of the ternary complex with GTP and Met-tRNA$_i^{Met}$ and subsequent formation of the ribosomal pre-initiation complex and/or joining of the large (prokaryotic 50S/ eukaryotic 60S) and small (prokaryotic 30S / eukaryotic 40S) subunits. The three-domain elongation factor EF-Tu/EF-1α form ternary complex with GTP and aminoacyl-tRNA and allows incorporation of the correct amino acid to a growing polypeptide chain in a translating ribosome. EF-G/EF-2 translocase has five domains and functions in moving the tRNA and mRNA down the ribosome.

These translation factors and the Ras-like small monomeric G-proteins have been extensively characterized and found to play critical roles thus being implicated in many physiological disorders; the OBG family, in contrast, has been poorly understood despite having an ancient lineage.

1.2 The Developmentally Regulated GTP binding proteins – an OBG subfamily member

The OBG family belonging to the OBG-HflX-like superfamily of the TRAFAC class of GTPases consists of five main subfamilies: Obg and YyaF/YchF found in eukaryotes and bacteria, DRG and NOG found in archaea and eukaryotes, and Ygr210 found in archaea and fungi (Leipe *et al.,* 2002). The Obg subfamily of the OBG class of proteins was found to be essential for vegetative growth and sporulation, while the NOG subfamily was involved in ribosome biogenesis and assembly. However, the roles of the YyaF/YchF, Ygr210 and DRG subfamilies are completely unknown. This study focuses on the DRG subfamily of proteins belonging to the OBG family of GTPases.

The **D**evelopmentally **R**egulated **G**TP binding proteins (DRGs) are strikingly conserved in archaea, fungi, plants and animals and are characterized by their abundant expression in growing cells and subsequent down-regulation in adult tissues. Their striking conservation throughout the major kingdoms suggested essential roles in fundamental pathways, but as of today the role of these proteins in the eukaryotic or archaebacterial cell is still a mystery. DRGs have two homologs, *drg1* and *drg2*, in most organisms from humans to plants sharing around 58% identity in human and mouse (O'Connell *et al.,* 2009). In archeaebacteria however, there seems to be only one *drg* gene, the product showing ~45% similarity to both human *drg1* and *drg2* (Li and Trueb 2000).

Drg1 protein encoded by the *drg1* gene was initially identified from the central nervous system of mouse in 1992 by a subtractive cloning approach (Kumar *et al.,* 1993, Sazuka *et al.,* 1992, Sazuka *et al.,* 1992). Initially named as NEDD-3 (Neural precursor cell expressed developmentally down-regulated protein 3) the expression of this gene was down-regulated in brain during development; being expressed in much higher levels in developing and differentiating brain than the fully differentiated adult brain. Hence it was speculated that they might play a role in early neurogenesis and renamed the NEDD-3 gene to *drg* and the protein that it encodes for, as DRG for **D**evelopmentally **R**egulated **G**TP binding protein. Data base searches revealed homologous sequences in *Drosophila* and *Halobacterium* and showed sequence similarities to a single biochemical class of proteins, the GTP binding proteins, with characteristic G1-G5 motifs for GDP/GTP exchange, GTP-induced conformational changes and GTP hydrolysis. Other than the G1-G5 region, little similarity was seen to other G-proteins except for the *Bacillus subtilis* Spo0B protein which is an Obg subfamily member (Trach and Hoch 1989).

Drg2 was first identified as a gene down regulated in transformed human fibroblasts and its mRNA was found to be mainly localised in skeletal muscle and heart (Schenker *et al.,* 1994) unlike Drg1 murine mRNA, which was abundant in the brain and not in heart. In the latter studies, Drg1 and Drg2 seemed to be localizing mainly in the cytoplasm (Sazuka

et al., 1992, Sommer *et al.,* 1994). It was also shown by immunofluorescence with myc-tagged Drg1 in *Candida albicans* that the protein localized in the cytoplasmic and perinuclear region (Chen and Kumamoto 2006).

The Ras and Rab related GTPases are generally farnesylated and geranylgeranylated at specific sequences located at the C-terminus like -C-A-A-X, -C-C or –C-X-C (A, nonpolar residue; X, any residue). These "prenylations" are typical of several of such low molecular weight GTPases and has been found to be essential for their membrane localization. Contrastingly, DRG and its homologues lacked these carboxy-terminal motifs for membrane anchoring, thus suggesting that the DRG proteins might be largely cytoplasmic.

Although consistently similar in most eukaryotic species, plant DRG proteins however have slightly different characteristics. In *Arabidopsis thaliana* and *Pisum sativum*, the presence of a third DRG member (Drg3) was detected in addition to Drg1 and Drg2. AtDrg1 and atDrg2 proteins were 56% identical, while atDrg2 and atDrg3 were found to be 95% identical to each other suggesting that the latter two would have been the result of a gene duplication event. Furthermore, both atDrg2 and atDrg3 contained 399 aminoacids unlike eukaryotic Drg1 proteins including atDrg1, which has generally around 369 amino acid residues (O'Connell *et al.,* 2009).

1.3 DRG proteins as GTPases

Initial descriptions of the DRG proteins had shown that *Drosophila melanogaster* and mouse DRG1 binds radiolabeled GTP *in vitro* (Sazuka *et al.,* 1992, Sommer *et al.,* 1994). Plant Drg1 (Pea) was also observed to bind to GTP-agarose which could be partly inhibited by GTP but not GDP or ATP (Nelson *et al.,* 2009). Despite this, detailed analysis of the binding kinetics or GTP hydrolytic activity of DRGs is still lacking. The Obg and YyaF/YchF subfamily members and *Arabidopsis thaliana* Drg1 are the only ones whose unique nucleotide binding and hydrolytic properties have been somewhat studied. They have been demonstrated to have a slow rate of GTP hydrolysis and rapidly exchange GDP for GTP in the absence of exogenous exchange factors unlike the heterotrimeric and Ras-like G-proteins.

1.4 DRGs and their binding partner - the DFRP proteins

Homo sapiens DRGs are highly homologous to its *Saccharomyces cerevisiae* counterparts, in that DRG1 shares around 66% identity and 80% similarity and DRG2 shares around 59% identity in amino acid sequence between the two species. In yeast, Drg1 and Drg2 are generally called in the literature as **R**ibosome **B**inding **G**TPase 1 (Rbg1) and **R**ibosome **B**inding **G**TPase 2 (Rbg2) respectively.

At about the same time as the identification of the presence of two homologous *drg* genes in each eukaryotic species (Li and Trueb 2000), it was discovered that yeast Rbg1 and Rbg2 forms complexes with the Gir2 (**G**enetically **I**nteracting with **R**ibosomal genes 2) protein (Ito *et al.,* 2001, Uetz *et al.,* 2000). This paved the way for the identification of the **D**RG **F**amily **R**egulatory **P**roteins (DFRPs) which were found to be required for maintaining normal levels of DRG proteins, with Dfrp1 binding specifically to Drg1 while Dfrp2 (Gir2 being Dfrp2) bound preferentially to Drg2 (Ishikawa *et al.,* 2005). Dfrp1 and Dfrp2 were also found to be highly conserved among eukaryotic species. Human Dfrp1 was known as the **L**ikely ortholog of mouse immediate **E**arly **R**esponse **E**rythro**PO**ietin 4 (Lerepo4) whilst the yeast Dfrp1 was found to be the **T**ranslation **M**achinery **A**ssociated protein 46 (Tma46). Human Lerepo4 and yeast Tma46 proteins are around 55% similar. The human counterpart of yeast Gir2 (Dfrp2) was called the **RWD D**omain-containing protein 1 (RWDD1) and share around 42% similarity. The human and yeast DRG-DFRP protein complexes taken up in this study are given in table 1.1.

Table 1.1: Human and yeast DRG-DFRP complexes studied.

Name of protein	Acronym	Source organism	No of aa	Uniprot id
Ribosome Binding GTPase 1	Rbg1	Yeast Drg1	369	P39729
Translation Machinery Associated protein 46	Tma46	Yeast Dfrp1	345	Q12000
Ribosome Binding GTPase 2	Rbg2	Yeast Drg1	368	P53295
Genetically Interacts with Ribosomal genes protein 2	Gir2	Yeast Dfrp1	265	Q03768
Developmentally Regulated GTP binding protein 1	Drg1	Human Drg1	367	Q9Y295
Likely ortholog of mouse immediate Early Response ErythroPOietin 4	Lerepo4	Human Dfrp1	426	Q8WU90

The function of the DFRP proteins in the cell has not been clearly characterized. As its name indicates, Lerepo4 was found under conditions suggestive of induction by erythropoietin (Gregory *et al.,* 2000). Its gene expression was also previously identified in a differential display screen to be modified by HIV-1 infection (Scheuring *et al.,* 1998). HIV-1 infection was shown to induce the gene expression of Lerepo4 and also, down-regulation of Lerepo4 expression by siRNAs reduced HIV-1 replication significantly, implying that Lerepo4 promoted HIV replication (Capalbo *et al.,* 2010). In yet another study, Lerepo4 was seen to be upregulated in rat neuronal cells upon stimulation by the nerve growth factor (NGF) suggesting a role in the regulation of neuronal survival and apoptosis (Heese *et al.,* 2004). Gir2, a highly acidic protein, had anomalous electrophoretic migrating behaviour on SDS-PAGE with an unexpectedly large Stokes radius in size exclusion chromatography. It is thermostable and highly sensitive to proteolysis. All these facts led to the identification that Gir2 might be an intrinsically unstructured protein (IUP) (Alves and Castilho 2005, Alves *et al.,* 2004).

1.5 DFRPs prevent degradation of DRGs by the ubiquitin-proteasome pathway

Ishikawa *et al.* have described that the complex formation between Drg1-Dfrp1 and Drg2-Dfrp2 would probably be occurring in the cytosol using immunofluorescence analysis of the subcellular localization of these proteins. The spatial and temporal expression patterns of *drg1* and *dfrp1* transcripts were also found to be similar in embryo and adult *Xenopus laevis*, being almost identical at some stages of embryonic development. Expression studies of the mouse DRG proteins on its own by transient transfection in 293T cells had resulted in inability to overexpress these proteins. This phenomenon was also previously reported for Drg1 (Mahajan *et al.,* 1996, Sazuka *et al.,* 1992). However, Drg1 and Drg2 expression dramatically increased when co-expressed with corresponding Dfrp1 and Dfrp2 proteins respectively. Dfrp1-deficient Chicken B cell lines also showed reduced Drg1 expression which could be restored using full length Dfrp1 expression vectors. The reduction of Drg1 was observed at the protein level and not in the mRNA level suggesting post-transcriptional regulation.

The Dfrp1 N-terminal zinc fingers might have the propensity to interact with ubiquitin and additionally the RWD domain of Dfrp2 is structurally related to the ubiquitin conjugating enzyme (E2) and E2 variant (UEV). The authors hence proposed that the increase in the stability of DRGs on physical association to DFRPs could possibly be by blocking ubiquitin-proteasome pathway degradation of the DRG proteins. Accordingly, it was demonstrated by immunoprecipitation assays that in the presence of a 26S proteasome inhibitor (MG132) an accumulation of polyubiquitinylated Drg1 and Drg2 occurred, which was progressively inhibited by increasing amounts of Dfrp1 and Dfrp2 respectively (Ishikawa *et al.,* 2003, Ishikawa *et al.,* 2005).

1.6 Functional characterization of the DRG proteins

Although many studies have implicated the DRG proteins to play a role in translation, differentiation and growth, the exact function of this category of TRAFAC class GTP-binding proteins has been difficult to understand.

Early studies have shown that Drg1 *in vitro* and *in vivo*, in both mouse and human (Mahajan *et al.,* 1996, Zhao and Aplan 1998) interacts with the oncogenic T-cell acute lymphoblastic leukemia (Tal1/Scl) protein, a basic helix-loop-helix (bHLH) transcription factor involved in cell growth and differentiation (hematopoietic development). Drg1 and another bHLH factor E47 bound to Tal1 in a mutually exclusive manner and competed to interact with Tal1. Drg1 could also bind to Tal2 and Lyl1 which are Tal1 related proteins also having the bHLH domain. It was speculated that in its GTP bound/unbound form Drg1 could bind to Tal1 or release it, which in turn could then translocate to nucleus and bind to its target genes. Mahajan *et. al.* also showed that overexpression of Drg1 increased rat

embryonic fibroblast transformation induced by c-myc and *ras* overexpression, affecting both the onset and size of foci formed. Drg2, as mentioned earlier, was also reported to be downregulated in SV-40 transformed (non-tumorigenic) fibroblasts in comparison to normal fibroblasts by a subtractive cDNA cloning approach (Schenker *et al.,* 1994).

In yet another study, filamentous invasion into agar matrices by *C. albicans* was attenuated by a Drg1 mutant, whereas it did not affect growth in liquid media or filamentous growth in serum. In the agar, both at 25ºC and 30ºC Drg1 null mutant showed defective growth which could be partially complemented by reintroduction of Drg1 gene at its native locus. These results suggested that Drg1 might be important for filamentous growth specifically in matrix-embedded conditions. The null mutant also was defective in filamentation when embedded in YPS, Spider or low ammonium minimal media at 25ºC and 30ºC. Intravenous injection of the *C. albicans* Drg1 null mutant into mice was also found to cause delayed lethality which in addition, exhibited reduced invasion into tissues (Chen and Kumamoto 2006).

Drg1 was also found to interact with Efg1, another bHLH transcription factor using the yeast two-hybrid system. C-terminal tagged Drg1 failed to interact with Efg1 implying that Drg1 might bind to bHLH transcription factors through its C-terminal domain as also suggested for Tal1. Efg1 is a repressor of embedded filamentation in *C. albicans* and it was suggested that Drg1 might play a role in regulating Efg1 as double deletions of *drg1* and *efg1* resulted in hyperfilamentation similar to a single *efg1* null mutant. That is to say that the reduced filamentation observed with the single *drg1* null mutant was not observed in the Δ*drg1*Δ*efg1* mutant (Chen and Kumamoto 2006).

In an Ubc9 fusion-dependent SUMOylation system (UFDS), Drg1 showed stimulation-dependent SUMOylation induced by the MEKK1 Map3 kinase (Jakobs *et al.,* 2007). The human protein kinase MPSK1 (Myristoylated and palmitoylated serine/threonine kinase 1) also known Stk16, was identified to be an interaction partner of Drg1 by yeast two hybrid and GST pull-down assays. The interaction was independent of the presence of nonhydrolyzable GDP or GTP and the region was mapped onto the N-terminal 1-65 residues of Drg1. MPSK1 was shown to phosphorylate Drg1 at Thr100 present in the G1 motif (Eswaran *et al.,* 2008). Drg1 was also seen to be present in the PI(4,5)P2 and PI(3,4,5)P3 interactomes of the LIM1215 colonic carcinoma cell line as according to an affinity based assay using phosphoinositides immobilized on beads or incorporated into liposomes followed by nanoRP-HPLC ESI MS/MS analysis (Catimel *et al.,* 2008, Catimel *et al.,* 2009).

1.6.1 Association of DRG factors to translation

The Drg1/Dfrp1 complex was found to be co-localizing with polysomes while the Drg2/Dfrp2 complex was found mostly in the non-polysomal fractions by sedimentation analyses using mouse liver homogenates (Ishikawa *et al.,* 2009). Plant DRG proteins were also shown to be capable of binding to ribosomes, a process partly inhibited by GTPγS (Nelson *et al.,* 2009). Indirect ribosome-Tma46-Drg1 (Fleischer *et al.,* 2006) and yeast Rbg1-Gir2-Gcn1-ribosome (Wout *et al.,* 2009) associations were also suggested. Consistent with these previously available information, it had recently been shown by our collaborators (Dr. Bertrand Seraphin, IGBMC, France), using gradient fractionation of extracts carrying epitope tagged versions of the proteins, that the yeast Drg1-Dfrp1 complex, namely Rbg1-Tma46, associates with the polysomes of the translation machinery (Daugeron *et al.,* 2011). Upon polysome disruption, the complex associates to the 80S monosomes and further with 40S and 60S ribosomal subunits. Rbg2-Gir2 on the other hand did not appear to associate with translating ribosomes and was suggested to be probably recruited to polysomes by Gcn1 under specific conditions. Gcn1 is a translational regulator involved in general amino acid control which was recently shown to interact with Gir2 (Wout *et al.,* 2009).

Deletion of the two yeast DRG-DFRP complexes alone or in combination does not result in a detectable phenotype, thus making it difficult to understand the function of these factors. Interestingly, Daugeron *et al.* recently reported also that a triple deletion mutant of Rbg1, Rbg2 (yeast Drg homologs) and Slh1 (Ski2-like helicase 1, a putative RNA helicase) exhibits a strong negative growth phenotype in yeast which was not observed with double deletants. Synthetic growth phenotypes were also observed for other combination of mutations inactivating simultaneously the two yeast DRG-DFRP complexes (namely Rbg1-Tma46, Rbg2-Gir2) and Slh1, demonstrating a functional redundancy of these factors. This observation provided for the first time the possibility to perform functional assays for Rbg proteins. Polysome analyses indicated that translation was impaired in the triple deletion mutant with profiles reminiscent of the pattern observed in translation initiation mutants such as cdc33-1 (eIF4E). The amount of polysomes decreased with a subsequent increase in the 80S peak in the mutant strains compared to the wild type strains. Growth complementation analyses using plasmid-borne epitope-tagged versions of the protein allowed the study of the contribution of some of the known conserved motifs in Rbg1 and Tma46, to the function of this complex. This demonstrated that the Rbg1 GTPase activity and Tma46 zinc fingers were required for its function, mutations in the G1 motif (for e.g.: $Rbg1_{GFPSVGKN}$) of Rbg1 and CCCH Zn fingers of Tma46 affecting growth as well as polysome association (Daugeron *et al.,* 2011).

1.7 Previous knowledge about structure

In addition to the characteristic GTP binding domain, DRGs (and YyaF/YchF subfamily) are also said to contain an N-terminal glycine-rich motif and a C-terminal TGS domain with unknown function. Among the OBG family members, only a few structures have been reported so far and they belong to the Obg and YyaF/YchF subfamily. The only structure reported for the Drg subfamily is a NMR solution structure of the TGS domain containing 93 amino acids (PDB id: 2EKI). This domain is shared by threonyl-tRNA synthetases (ThrRSs), the DRG subfamily (GTP-binding proteins) and guanosine polyphosphate phosphohydrolases/ synthetases (SpoT/RelA) (Wolf *et al.,* 1999). The TGS domain is speculated to be a RNA binding domain but its function is as yet not been experimentally demonstrated.

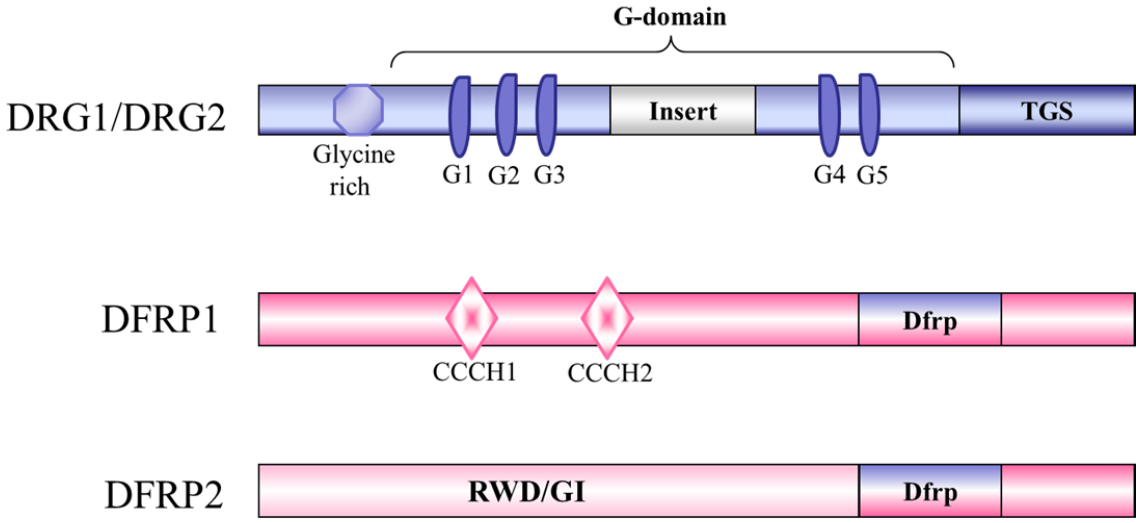


Figure 1.2: Schematic drawing of the previously known domain organization of DRG and DFRP proteins. DRG1 and DRG2 being highly homologous to each other, a single representation is given depicting the glycine rich region at the N-terminus, the GTP binding motifs (G1-G5), the insert sequence between G3 and G4 and finally the TGS domain at the C-terminus. The general representation of the domain organization of the DFRP1 and DFRP2 proteins is shown with the C-terminal region (Dfrp) that interacts with DRGs. Also shown are the CCCH type zinc fingers of DFRP1 and the RWD domain of DFRP2.

Drg1 and Drg2, as mentioned earlier, are highly homologous to each other, but this is not the case with Dfrp1 and Dfrp2 proteins. Apart from the C-terminal region containing the dfrp domain which was identified to be the minimum region necessary for binding to DRGs (Ishikawa *et al.,* 2005), DFRP proteins are very different in the N-terminal region. Dfrp1 proteins (Tma46 and Lerepo4) comprise of two characteristic CCCH-type zinc fingers at their N-terminal. On the other hand, Dfrp2 proteins (Gir2 and Rwdd1) contain an N-terminal RWD domain (also called GI domain, (Kubota *et al.,* 2000) present in **R**ING finger-containing proteins, **W**D-repeat-containing proteins, and yeast **D**EAD (DEXD)-like

helicases (Figure 1.2). Yeast Gir2 interacts with the Gcn1 protein involved in general amino acid control pathway through its RWD domain (Wout *et al.,* 2009). The RWD domain structure of another protein, Gcn2 which also interacts with Gcn1 has been solved. However, structural knowledge about the DFRP proteins especially the dfrp domain is not yet available.

1. Introducción

1.1 La superfamilia GTPasa

Las proteínas de unión a GTP o proteínas G representan una categoría diversa de moléculas reguladoras en la célula y son componentes de las rutas de transducción de señales en eucariotas y, de la misma manera, en procariotas. Las proteínas G se caracterizan por su capacidad común de unión e hidrólisis de guanosina 5'-trifosfato (GTP) a guanosina 5'-difosfato (GDP), actuando como interruptores moleculares binarios entre su forma unida a GTP (activa) y a GDP (inactiva). El dominio de unión a GTP (dominio G) de estas proteínas contiene cinco motivos característicos llamados G1 (Walker A/P-loop), G2, (Switch I), G3 (Walker B/Switch II), G4 and G5. El motivo G1 (GxxxxGK(S/T)) es responsable de la unión de grupos α- y β-fosfato del nucleótido, el motivo G2 (x(T/S)x) está implicado en la unión a Mg^{2+}, mientras que el G3 (DxxG) se une tanto a Mg^{2+} como al γ-fosfato. G4 ((N/T)KxD) se une directamente con el nucleótido a través de los residuos K y D y el G5, débilmente conservado, actúa principalmente en el reconocimiento de bases de guanina. Las GTPasas P-loop, llamadas así por su motivo Walker A entre una hélice y una hoja (Walker et al., 1982), funcionan produciendo cambios conformacionales entre las formas unidas a GTP/GDP de las señales de transducción celular de la proteína G a proteínas efectoras aguas abajo principalmente a través de cambios en las regiones de los interruptores I (G2) y II (G3). (Bourne *et al*., 1990, Bourne *et al*., 1991, Wittinghofer y Vetter, 2011).

Las enzimas que utilizaban GTP en lugar de ATP se caracterizaron en primer lugar como factores de iniciación, elongación y terminación en el proceso de la síntesis de proteína en la célula; siendo el factor de elongación Tu (EF-Tu) la primera estructura tridimensional obtenida de una GTPasa (Kabsch *et al*., 1977). A continuación, se identificó la familia de proteínas oncogénicas Ras implicadas en varios procesos como expresión de genes (Ras, Rho), reorganización del citoesqueleto (Rho, Rac), importación nuclear (ran), transporte vesicular (Rab, Sar1/Arf) y mitogénesis (Ras, Rap) (García-Ranea y Valencia, 1998, Macara *et al*., 1996). La comparación de estudios estructurales y genómico-secuenciales han identificado numerosas GTPasas que juegan un papel en diversos procesos biológicos/celulares, regulando la forma activa un amplio espectro de procesos celulares. De acuerdo con esto, la superclase de GTPasas se ha dividido en dos clases, la clase TRAFAC (del inglés, *TRAnslation FACtor-related*) que contiene a la mayoría de GTPasas y la clase SIMIBI (del inglés, *Signal recognition GTPases, MinD and BioD superfamily*) (Leipe *et al*., 2002, Verstraeten *et al*., 2011). Las subsecuentes divisiones de la clase TRAFAC se han representado en la figura 1.1.

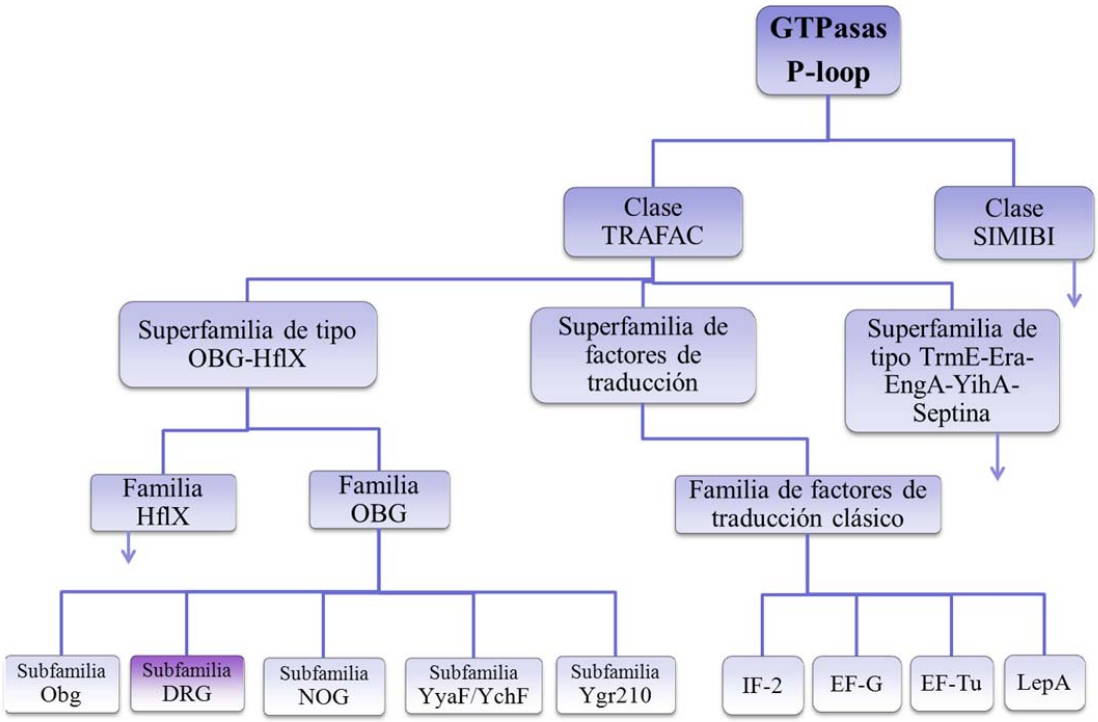


Figura 1.1. Clasificación de las P-loop GTPasas. El dendrograma de clasificación de las familias de GTPasas de la clase TRAFAC se ha representado destacando la familia OBG, sin incluir algunas de sus ramificaciones, que no son relevantes para este estudio. La figura ha sido modificada de Verstraeten *et al*., 2011.

La familia de factores de traducción clásicos está extendida en los tres superreinos (Archaea, Bacteria y Eucariotas) y podría ser una herencia vertical del último ancestro común universal (LUCA, del inglés *Last Universal Common Ancestor*) de todas las formas de vida. Éstos incluyen GTPasas implicadas en la maquinaria de traducción. Los factores de iniciación GTPasas están involucrados en la formación del complejo ternario con GTP y Met-tARN$_i^{Met}$ y la formación subsecuente del complejo de preiniciación ribosomal y/o la unión de las subunidades grande (50S procariótico/ 60S eucariótico) y pequeña (30S procariótico/ 40S eucariótico). El factor de elongación EF-Tu/Ef-1α de tres dominios forma complejos ternarios con GTP y aminoacil-ARNt y permite la incorporación del aminoácido correcto a una cadena polipeptídica en construcción en el ribosoma durante la traducción. La translocasa EF-G/EF-2 tiene cinco dominios y funciona desplazando el ARNt y el ARNm hacia abajo en el ribosoma.

Estos factores de traducción y las proteínas G monoméricas pequeñas de tipo Ras se han caracterizado extensamente y se ha encontrado que juegan papeles críticos, estando implicadas en muchos desórdenes fisiológicos; por el contrario, a pesar de que la familia OBG tiene un antiguo linaje, es poco conocida.

1.2 Las proteínas de unión a GTP reguladas en el desarrollo – un miembro de la subfamilia OBG

La familia OBG, que pertenece a la superfamilia OBG-HflX de la clase TRAFAC de GTPasas, consta de cinco subfamilias principales: Obg y YyaF/YchF en eucariotas y bacterias, DRG y NOG en archaea y eucariotas e Ygr210 en archaea y hongos (Leipe *et al*., 2002). Se sabe que la subfamilia Obg de la clase OBG de proteínas es esencial para el crecimiento vegetativo y la esporulación, mientras que la subfamilia NOG está implicada en la biogénesis y ensamblaje del ribosoma. Sin embargo, las funciones de las subfamilias YyaF/YchF, Ygr210 y DRG son completamente desconocidas. Este estudio se centra en la subfamilia de proteínas DRG pertenecientes a la familia OBG de GTPasas.

Las proteínas de unión a GTP reguladas en el desarrollo (DRGs) están notablemente conservadas en archaea, hongos, plantas y animales y se caracterizan por su abundante expresión en células en crecimiento y en la subsecuente regulación negativa en tejidos adultos. Esta conservación presente en los principales reinos sugería papeles esenciales en rutas fundamentales, pero la función de estas proteínas en células eucarióticas o de arqueobacterias continúa siendo un misterio hasta la fecha. DRG tiene dos homólogos, *drg1* y *drg2*, en la mayor parte de los organismos desde humanos hasta plantas, que comparten cerca del 58% de identidad entre humano y ratón (O'Connell *et al*., 2009). Sin embargo, en arqueobacterias parece que sólo está presente un gen *drg*, cuyo producto muestra un ~45% de similitud con los *drg1* y *drg2* humanos (Li y Trueb, 2000).

La proteína Drg1 codificada por el gen *drg1* se identificó inicialmente en el sistema nervioso central de ratón en 1992 en una aproximación de clonaje substractivo (Kumar *et al*., 1993, Sazuka *et al*., 1992, Sazuka *et al.*, 1992). La expresión de este gen, inicialmente llamado NEDD-3 (proteína 3 regulada negativamente en el desarrollo, que se expresa en células neuronales precursoras), estaba regulada negativamente en el cerebro durante el desarrollo; expresándose en niveles mucho mayores durante el desarrollo y diferenciación del cerebro que en el cerebro adulto totalmente diferenciado. Por ello, se especuló que podría tener una función en la neurogénesis temprana y se renombró el gen NEDD-3 como *drg* y la proteína para la que codifica como DRG (proteína de unión a GTP regulada en el desarrollo). Las búsquedas en bases de datos revelaron secuencias homólogas en *Drosophila* y *Halobacterium* y mostraron similitud de secuencia con una única clase bioquímica de proteína, las proteínas de unión a GTP, con motivos característicos G1-G5 para el intercambio GDP/GTP, cambios conformacionales inducidos e hidrólisis de GTP. Aparte de las regiones G1-G5, no se observó similitud con otras proteínas G a excepción de la proteína Spo0B de *Bacillus subtilis*, que es un miembro de la subfamilia Obg (Trach y Hoch, 1989).

Drg2 se identificó por primera vez como un gen regulado negativamente en fibroblastos humanos transformados y se encontró que su ARNm está localizado principalmente en músculo esquelético y el corazón (Schenker *et al.*, 1994), a diferencia del ARNm murino de Drg1, que resultó ser abundante en el cerebro pero no en el corazón. En los últimos estudios, Drg1 y Drg2 parecían estar localizados principalmente en el citoplasma (Sazuka *et al*,. 1992, Sommer *et al.*1994). Asímismo, se comprobó por inmunofluorescencia con Drg1 etiquetada con myc en *Candida albicans* que la proteína estaba localizada en la región citoplasmática y perinuclear (Chen y Kumamoto, 2006).

Las GTPasas relacionadas con Ras y Rab están generalmente farnesiladas y geranilgeraniladas en secuencias específicas localizadas en el C-terminal, como -C-A-A-X, -C-C, o bien –C-X-C (A, residuo no polar; X, cualquier residuo). Estas prenilaciones son típicas de varias GTPasas de bajo peso molecular y se ha visto que son esenciales para su localización en la membrana. Sin embargo, DRG y sus homólogos carecen de estos motivos carboxiterminales para el anclaje a la membrana, sugiriendo con esto que las proteínas DRG podrían ser mayoritariamente citoplasmáticas.

Aunque son similares en la mayoría de especies eucariotas, las proteínas DRG de plantas presentan características ligeramente diferentes. En *Arabidopsis thaliana* y *Pisum sativum*, se detectó la presencia de un tercer miembro DRG (Drg3), además de Drg1 y Drg2. Las proteínas atDrg1 y atDrg2 presentaban un 56% de identidad, mientras que atDrg2 y atDrg3 mostraban un 95% de identidad entre sí, sugiriendo que estas últimas podrían ser el resultado de un evento de duplicación de genes. Además, atDrg2 y atDrg3 contenían 399 aminoácidos, a diferencia de las proteínas Drg1 eucarióticas, incluyendo a atDrg1, que normalmente tiene cerca de 369 aminoácidos (O'Connell *et al*., 2009).

1.3 Las proteínas DRG como GTPasas

Las descripciones iniciales de las proteínas DRG habían mostrado que DRG1 de *Drosophila melanogaster* y de ratón se unían *in vitro* a GTP marcado radiactivamente (Sazuka *et al*., 1992, Sommer *et al*., 1994). Se observó también que Drg1 de plantas (guisante) se unía a GTP-agarosa que podría estar parcialmente inhibida por GTP pero no por GDP o ATP (Nelson *et al.*, 2009). A pesar de esto, el análisis detallado de las cinéticas de unión o de la actividad hidrolítica de GTP de las DRGs todavía se desconoce. Los miembros de la subfamilia Obg e YyaF/YchF y Drg1 de *Arabidopsis thaliana* son las únicas cuyas propiedades singulares de unión e hidrólisis de nucleótidos se han estudiado en cierta manera. Éstas han demostrado tener una baja tasa de hidrólisis de GTP e intercambio rápido de GDP a GTP en ausencia de factores de intercambio exógenos a diferencia de las proteínas G heterotriméricas y de tipo Ras.

1.4 DRGs y su ligando celular – las proteínas DFRP

Las DRGs de *Homo sapiens* presentan alta similitud con sus homólogos de *Saccharomyces cerevisiae*, de tal forma que DRG1 comparte en torno al 66% de identidad y un 80% de similitud y DRG2 comparte un 59% de identidad en la secuencia de aminoácidos entre ambas especies. En levadura, Drg1 y Drg2 se denominan generalmente en la literatura como Rbg1 (del inglés, *Ribosome Binding GTPase 1*) y Rbg2 respectivamente.

Al mismo tiempo que se identificó la presencia de los dos genes homológos *drg* en cada especie eucarióticа (Li y Trueb, 2000), se descubrió que las proteínas de levaduras Rbg1 y Rbg2 forman complejo con la proteína Gir2 (*Genetically Interacting with Ribosomal genes 2*) (Ito *et al.*, 2001, Uetz *et al*., 2000). Esto allanó el camino para la identificación de la familia reguladora de proteínas de DRG (DFRP), que se requería para mantener niveles normales de las proteínas DRG, mediante la unión específica de Dfrp1 a Drg1, mientras que Dfrp2 (siendo Gir2 Dfrp2) se unía preferentemente a Drg2 (Ishikawa *et al*., 2005). Además, se observó que Dfrp1 y Dfrp2 se encontraban altamente conservadas entre especies eucarióticas. Dfrp1 humana se conocía como Lerepo4 (*Likely ortholog of mouse immediate early response erythropoietin 4*) mientras que Dfrp1 de levadura se denominó Tma46 (*Translation machinery associated protein 46*). Las proteínas Lerepo4 humana y Tma46 de levadura presentan una similitud del 55%. El homólogo humano de levadura Gir2 (Dfrp2) se nombró RWDD1 (*RWD domain-containing protein 1*) y comparten en torno al 42% de similitud. Los dos complejos de proteínas DRG-DRFP de ambos organismos, humano y levadura que están estudiado en este trabajo se muestran en la tabla 1.1.

Table 1.1: Complejos DRG-DFRP estudiados en humano y levadura.

Nombre de la proteína	Acrónimo	Organismo	Nº de aá	Uniprot id
Ribosome Binding GTPase 1	Rbg1	Drg1 (levadura)	369	P39729
Translation Machinery Associated protein 46	Tma46	Dfrp1 (levadura)	345	Q12000
Ribosome Binding GTPase 2	Rbg2	Drg1 (levadura)	368	P53295
Genetically Interacts with Ribosomal genes protein 2	Gir2	Dfrp1 (levadura)	265	Q03768
Developmentally Regulated GTP binding protein 1	Drg1	Drg1 (humano)	367	Q9Y295
Likely ortholog of mouse immediate Early Response ErythroPOietin 4	Lerepo4	Dfrp1 (humano)	426	Q8WU90

Aun no se ha podido elucidar claramente la función de la proteína DFRP en la célula. Como su nombre indica, Lerepo4 fue encontrada bajo condiciones que indicaban inducción por eritropoyetina (Gregory *et al.*, 2000). Al mismo tiempo se vio que la expresión diferencial del gen quedaba alterada con la infección por VIH-1. Se demostró que el VIH-1 al infectar la célula era capaz de inducir la expresión de Lerepo4 y, a su vez, la regulación negativa de la expresión de Lerepo4 mediante siRNA reducía significativamente la tasa de replicación del virus lo que demostró el papel de Lerepo4 a la hora de activar la replicación del VIH (Capalbo *et al.,* 2010). En otro estudio se vio que en las células nerviosas de rata Lerepo4 se encontraba sometido a un proceso de regulación positiva inducido por el factor de crecimiento nervioso (NGF) sugiriendo un papel en la regulación de la supervivencia neuronal y apoptosis (Heese *et al.,* 2004). Gir2, una proteína fuertemente ácida, presenta un patrón de migración anómalo en el proceso de electroforesis SDS-PAGE y al mismo tiempo muestra un radio de Stokes sorprendentemente grande cuando se somete a la cromatografía de exclusión molecular. Al mismo tiempo es termoestable y altamente sensible a la proteólisis. Todos esos factores llevaron a pensar que Gir2 podría ser una proteína intrínsecamente desestructurada (IUP) (Alves y Castilho 2005, Alves *et al.,* 2004).

1.5 Los DFRPs evitan la degradación de DRGs por la ruta de ubiquitín-proteasoma

Ishikawa *et al*. han descrito que la formación del complejo entre Drg1-Dfrp1 y Drg2-Dfrp2 podría ocurrir probablemente en el citosol, utilizando análisis de inmunofluorescencia de la localización subcelular de estas proteínas. Los patrones de expresión espacial y temporal de los transcritos de *drg1* y *dfrp1* eran similares en embriones y adultos de *Xenopus laevis*, siendo casi idénticos en algunos estadíos del desarrollo embrionario.

Los estudios de expresión de las proteínas DRG de ratón solas por transfección transitoria en células 293T resultaron en la incapacidad de sobreexpresar estas proteínas. Este fenómeno había sido descrito previamente para Drg1 (Sazuka *et al.*, 1992, Mahajan *et al*, 1996). Sin embargo, la expresión de Drg1 y Drg2 aumentó notablemente cuando se coexpresó con las correspondientes Dfrp1 y Dfrp2 respectivamente. Las líneas celulares B deficientes en Dfrp1 de pollo también mostraron una expresión reducida de Drg1, que podría ser restaurada utilizando la expresión de Dfrp1 completa utilizando vectores. La reducción de Drg1 se observó a nivel de proteína y no a nivel de ARNm, sugiriendo la existencia de una regulación posttranscripcional.

Los dedos de zinc N-terminales de Dfrp1 podrían ser propensos a interaccionar con ubiquitina y, además, el dominio RWD de Dfrp2 está relacionado estructuralmente con la enzima de conjugación de ubiquitina (E2) y la variante E2 (UEV). Por consiguiente, los autores propusieron que el aumento en la estabilidad de DRGs en la asociación física a

DFRPs podría producirse posiblemente por el bloqueo de la ruta de degradación de proteínas DRG por el ubiquitín-proteasoma. De acuerdo con esto, se ha demostrado por ensayos de inmunoprecipitación que en presencia de un inhibidor de proteasoma 26S (MG132), se producía una acumulación de Drg1 y Drg2 poliubiquitiniladas, que se inhibía progresivamente por cantidades crecientes de Dfrp1 y Dfrp2 respectivamente (Ishikawa *et al*., 2003, Ishikawa *et al*., 2005).

1.6 Caracterización funcional de las proteínas DRG

Aunque muchos estudios han implicado a las proteínas DRG como participantes en traducción, diferenciación y crecimiento, la función exacta de esta categoría de proteínas de unión a GTP de la clase TRAFAC es difícil de dilucidar.

Los estudios iniciales mostraron que Drg1 *in vitro* e *in vivo*, tanto en ratón como en humano, (Mahajan *et al*., 1996, Zhao y Aplan, 1998) interacciona con la proteína oncogénica de leucemia linfoblástica aguda de las células T (Tal1/Scl), un factor de transcripción bHLH (hélice-lazo-hélice básico) implicado en el crecimiento y diferenciación celular (desarrollo hematopoyético). Drg1 y otro factor bHLH E47 unido a Tal1 en una forma exclusiva para ambos, que además competían para interaccionar con Tal1. Drg1 podría también unirse a Tal2 y Lyl1, que son proteínas relacionadas con Tal1 que también contienen el dominio bHLH. Se especuló que Drg1 en su forma unida o separada de GTP se unía a Tal1 o liberaba Tal1 que a continuación podía translocarse al núcleo y unirse a sus genes diana. Mahajan *et al*. también comprobaron que la sobreexpresión de Drg1 aumentaba la transformación de fibroblastos embriónicos en rata inducidos por la sobreexpresión de c-myc y ras, afectando tanto al inicio como al tamaño de los focos de células. Drg2, como se mencionó anteriormente, se encontró por una aproximación de clonación substractiva de ADNc que estaba regulado negativamente en fibroblastos transformados SV-40 (no tumorigénicos), en comparación con fibroblastos normales (Schenker *et al*., 1994).

En un estudio posterior, la invasión filamentosa de *Candida albicans* en matrices de agar se atenuó por un mutante de Drg1, mientras que éste no afectó al crecimiento en medio líquido ni al crecimiento filamentoso en suero. En el agar, tanto a 25ºC como a 30ºC, el mutante nulo de Drg1 mostró un crecimiento defectivo que podría ser complementado parcialmente por la reintroducción del gen de Drg1 en su locus nativo. Estos resultados sugirieron que Drg1 podría ser importante para el crecimiento filamentoso específicamente en condiciones de embebido en la matriz. El mutante nulo también era defectivo en filamentación cuando se crecía en YPS, Spider o un medio mínimo bajo en amonio a 25ºC y 30ºC. La inyección intravenosa en ratones del mutante nulo de Drg1 de *C.albicans* se vio que causaba un retraso en la letalidad que, además, presentaba una invasión reducida en los tejidos (Chen y Kumamoto, 2006).

Utilizando el sistema de doble híbrido en levaduras, se observó también que Drg1 interaccionaba con Efg1, otro factor de transcripción bHLH. Drg1 etiquetado en el C-terminal no interaccionaba con Efg1, implicando que Drg1 podría unirse a factores de transcripción bHLH a través de su dominio C-terminal, también sugerido por la proteína Tal1. Efg1 es un represor de la filamentación embebida en *C. albicans* y se propuso que Drg1 podría tener un papel en la regulación de Efg1, puesto que la doble delección de *drg1* y *efg1* resultó en una hiperfilamentación similar a la que mostraba el mutante nulo sólo de *efg1*; es decir, la filamentación reducida que se observó con el mutante nulo único de *drg1* no se observó en el mutante Δdrg1Δefg1 (Chen y Kumamoto, 2006).

En un sistema de SUMOilación dependiente de fusión Ubc9 (UFDS), Drg1 mostró una SUMOilación dependiente de estímulo inducida por la kinasa Map3 MEKK1 (Jakobs et al., 2007). La proteína quinasa humana MPSK1 (Myristoylated and palmitoylated serine/threonine kinase 1), también conocida como Stk16, fue identificada como proteína que interaccionaba con Drg1 mediante ensayos de doble híbrido en levadura y por experimentos de pull-down. La interacción fue independiente de la presencia del GTP o de la forma no hidrolizable del GDP y la región responsable de la interacción se acotó a los residuos 1-65 de Drg1. Se demostró que la MPSK1 fosforila a Drg1 en el residuo de Thr100 ubicado en el motivo G1 (Eswaran *et al.,* 2008). Por otro lado, se vio que Drg1 formaba parte de los interactomas PI(4,5)P2 y PI(3,4,5)P3 de la línea celular del carcinoma de colon LIM1215. En este caso se utilizó un ensayo de afinidad empleando resinas de fosofinosítido inmovilizado o incorporado en los liposomas, seguido por los análisis mediante nanoRP-HPLC ESI MS/MS. (Catimel *et al.,* 2008, Catimel *et al.,* 2009).

1.6.1 Asociación de los factores DRG a la traducción

Se describió que el complejo Drg1/Dfrp1 colocalizaba con polisomas, mientras que el complejo Drg2/Dfrp2 se encontraba en la mayoría de los casos en las fracciones no polisomales por análisis de sedimentación utilizando homogeneizados de hígado de ratón (Ishikawa *et al*., 2009). Las proteínas DRG de plantas mostraron ser capaces de unirse a ribosomas, un proceso inhibido en parte por GTPγS (Nelson *et al*., 2009). Las asociaciones indirectas ribosoma–Tma46–Drg1 (Fleischer *et al*., 2006) y Rbg1–Gir2–Gcn1–ribosoma de levadura (Wout *et al*., 2009) se postularon del mismo modo. En consonancia con esta información, nuestros colaboradores (Dr. Bertrand Seraphin, IGBMC, France) han mostrado recientemente, utilizando fraccionamiento en gradiente de extractos que contenían versiones de las proteínas etiquetadas con epítopos, que el complejo de levadura Drg1-Dfrp1, llamado Rbg1-Tma46, se asocia con los polisomas de la maquinaria de traducción (Daugeron *et al*., 2011). Hasta la disociación del polisoma, el complejo se asocia a los monosomas 80S y posteriormente con las subunidades ribosomales 40S y 60S. Por el contrario, Rbg2-Gir2 no parecían asociarse con los ribosomas durante la traducción y se sugirió que probablemente eran reclutados en los polisomas por Gcn1 en condiciones

específicas. Gcn1 es un regulador traduccional implicado en el control general de aminoácidos que recientemente se ha demostrado que interacciona con Gir2 (Wout *et al*., 2009).

La delección de los dos complejos DRG-DFRP solos o en combinación no resultó en un fenotipo detectable, lo que hace más difícil el entendimiento de la función de estos factores. Es interesante destacar que Daugeron *et al*. recientemente describió que el triple mutante de la delección de Rbg1, Rbg2 (homólogos de levadura de Drg) y Slh1 (helicasa 1 de tipo Ski2, una ARN helicasa putativa) exhibe un evidente fenotipo de crecimiento negativo en levaduras, que no se observó en los dobles mutantes deleccionados. Los fenotipos sintéticos de crecimiento se estudiaron también para otras combinaciones de mutaciones, inactivando simultáneamente los dos complejos de levadura DRG-DFRP (denominados Rbg1-Tma46, Rbg2-Gir2) y Slh1, demostrando una redundancia funcional de estos factores. Esta observación proporcionó por primera vez la posibilidad de llevar a cabo análisis funcionales para las proteínas Rbg. Los análisis de polisomas indicaron que la traducción se vio perjudicada en el mutante de la triple delección con perfiles reminiscentes del patrón observado en los mutantes de iniciación de la traducción como cdc33-1 (eIF4E). La cantidad de polisomas disminuyó con el subsecuente aumento en el pico de 80S en las cepas mutantes respecto a las cepas de tipo silvestre. Los análisis de complementación de crecimiento utilizando versiones de la proteína etiquetadas con epítopos en plásmidos permitieron el estudio de la contribución de algunos de los motivos conservados conocidos en Rbg1 y Tma46 para la función de este complejo. Esto demostró que la actividad GTPasa de Rbg1 y los dedos de zinc de Tma46 eran necesarios para su función, viéndose que las mutaciones en el motivo G1 (ej.: $Rbg1_{GFPSVGKN}$) de Rbg1 y los dedos de zinc CCCH de Tma46 afectaban al crecimiento y a la asociación de polisomas (Daugeron *et al*., 2011).

1.7 Conocimientos previos sobre la estructura

Además del dominio de unión a GTP característico, se sabe que DRG (y la subfamilia YyaF/YchF) contiene un motivo N-terminal rico en glicina y un dominio TGS C-terminal de función desconocida. Entre los miembros de la familia OBG, sólo se han publicado unas pocas estructuras y éstas pertenecen a la subfamilia Obg e YyaF/YchF. La única estructura publicada para la subfamilia Drg es una estructura de RMN del dominio TGS que consta de 93 aminoácidos (PDB id: 2EKI). Este dominio es compartido por las sintetasas treonil-ARNt (ThrRSs), la subfamilia DRG (proteínas de unión a GTP) y las fosfohidrolasas / sintetasas de guanosina polifosfato (SpoT/RelA) (Wolf *et al*., 1999). El dominio TGS se especula que podría ser un dominio de unión a ARN pero su función no se ha demostrado experimentalmente todavía.

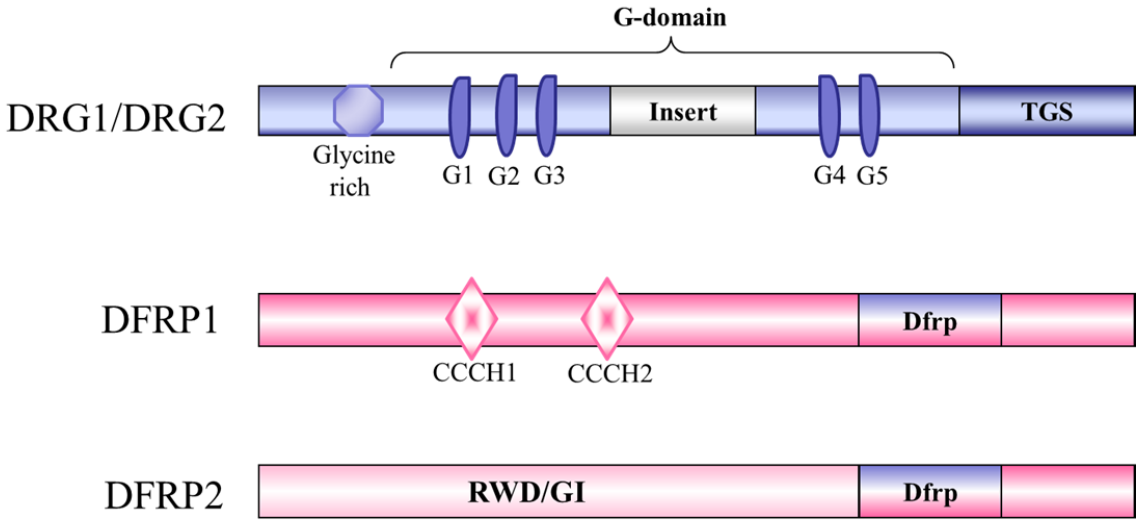


Figura 1.2: Esquema de la organización de dominios de las proteínas DRG y DFRP basada en conocimientos previos. Teniendo en cuenta la alta homología entre DRG1 y DRG2 se muestra una única figura que abarca la región N terminal rica en glicina, los motivos de unión al GTP (G1-G5), la secuencia insertada entre G3 y G4 y por último el dominio TGS en el extremo carboxi-terminal. El esquema general de la disposición de los dominios de las proteínas DFRP1 y DFRP2 incluye la región C-terminal (Dfrp) que interacciona con DRGs. También se han representado los dedos de zinc de tipo CCCH de DFRP1 y el dominio RWD de DFRP2.

Drg1 y Drg2, como se mencionó anteriormente, son altamente homólogos entre sí, pero esto no ocurre entre las proteínas Dfrp1 y Dfrp2. Aparte de la región C-terminal que contiene el dominio dfrp, que fue identificado como la región mínima necesaria para la unión a DRGs (Ishikawa *et al.,* 2005), estas proteínas DFRP son muy diferentes en la región N-terminal. Las proteínas Dfrp1 (Tma46 y Lerepo4) presentan dos dedos de zinc característicos de tipo CCCH en el N-terminal. Por el contrario, las proteínas Dfrp2 (Gir2 y Rwdd1) contienen un dominio RWD N-terminal (también llamado dominio GI, (Kubota *et al*., 2000) presente en proteínas que contienen dedos RING, proteínas que contienen repeticiones WD, y helicasas de levaduras de tipo DEAD (DEXD) (Figura 1.2). La proteína Gir2 de levadura interacciona con Gcn1, una proteína implicada en la ruta de control general de aminoácidos a través de su dominio RWD (Wout *et al.,* 2009). Se ha podido resolver la estructura del dominio RWD de Gcn2, otra proteína que también interacciona con Gcn1. Sin embargo, aun no existen conocimientos disponibles acerca de la estructura de las proteínas DFRP, sobre todo de su dominio Dfrp.

Objectives

2. Objectives

The Developmentally Regulated GTP binding proteins (DRGs) are found highly conserved in the two major kingdoms of life, the archaea and the eukaryota. They belong to the OBG family of proteins which form an ancient lineage in the TRAFAC class of GTPases. Drg1 and Drg2, which are the two paralogues of DRG present universally in most organisms, were shown to be regulated by the DFRP proteins, Dfrp1 and Dfrp2. Concomitantly, the Drg1-Dfrp1/Drg2-Dfrp2 complexes belonging to higher organisms are highly homologues to its yeast counterparts Rbg1-Tma46/Rbg2-Gir2 respectively, none of which have been structurally analysed. The integrated 3D Repertoire project by the European commission under the 6th Framework programme for solving protein complex structures from *S. cerevisiae* provided a perfect background for the study of these proteins.

Despite its high abundance and conservation, the exact role of the DRG proteins remains to be narrowed down. However, during the span of over 22 years of research on this subfamily, they have been largely implicated in cell growth and development, with afiliations to protein translation. Being overly expressed in growing and differentiating cells could provide direct implications of the DRG proteins towards cancer cell development and treatment for cancer.

With these premises, the following objectives were set up for this thesis doctoral work:

1. Determine the three-dimensional structure of the DRG protein in complex with its regulatory protein, DFRP by X-ray crystallography technique.
2. Characterise the interaction between DRG and DFRP and implications in its possible association to the translation machinery process.
3. Biochemically analyse the guanine nucleotide binding and hydrolytic properties of the DRG GTPases in the context of its component domains and complex formation with DFRP proteins.
4. Dissect the RNA binding properties of the DRG-DFRP complex proteins.

Materials and Methods

3. Materials & Methods

3.1 Production of the DRG-DFRP complex proteins in *E. coli* expression system

3.1.1 Obtainment of plasmids

The constructs for the three different complexes, yeast Rbg1-Tma46/Rbg2-Gir2 and human Drg1-Lerepo4 including the individual proteins, truncated and mutant forms used in this study are as given in table 3.1.

Table 3.1: DRG-DFRP plasmids for the *E. coli* expression system.

	Rbg1-Tma46	**Drg1-Lerepo4**	**Rbg2-Gir2**
DRG_{fl}-$DFRP_{dfrp}$	$_{6His}Rbg1_{fl}$-$Tma46_{205-345}$* $_{6His}Rbg1_{fl}$-$Tma46_{154-345}$* $_{6His}Rbg1_{fl}$-$Tma46_{137-345}$*		$_{6His}Rbg2_{fl}$-$Gir2_{174-238}$*
DRG_{fl}	$_{6His}Rbg1_{fl}$*	$_{6His}Drg1_{fl}$[a]	
$DFRP_{fl}$/ $DFRP_{dfrp}$		$_{6His}Lerepo4_{220-426}$[b] $_{6His}Lerepo4_{1-220}$[b]	$_{6His}Gir2_{fl}$*
$DRG_{\Delta tgs}$-$DFRP_{dfrp}$	$_{6His}Rbg1_{1-294}$-$Tma46_{205-345}$*		$_{6His}Rbg2_{1-294}$-$Gir2_{174-238}$*
DRG_{tgs}-$DFRP_{dfrp}$	$_{6His}Rbg1_{272-369}$-$Tma46_{205-345}$* $_{6His}Rbg1_{272-369}$-$Tma46_{154-345}$*		
DRG_{tgs}		$_{6His}Drg1_{289-367}$[a]	
$DRG_{\Delta s52dl}$-$DFRP_{dfrp}$	$_{6His}Rbg1_{\Delta 175-243+G}$-$Tma46_{205-345}$*		
DRG_{s5d2l}	$_{6His}Rbg1_{176-240}$	$_{6His}Drg1_{175-238}$	$_{6His}Rbg2_{174-240}$
	Rbg1-Tma46 mutants		
$DRG_{fl\ Inactive}$	$_{6His}Rbg1_{fl\ S79N}$-$Tma46_{205-345}$*		
$DRG_{fl\ Inactive}$	$_{6His}Rbg1_{fl\ VGKS\rightarrow VAMN}$-$Tma46_{205-345}$		
DRG_{fl}-$DFRP_{dfrp\ G272A\ R273E}$	$_{6His}Rbg1_{fl}$-$Tma46_{205-345\ G272A\ R273E}$*		
DRG_{fl}-$DFRP_{dfrp\ I241A\ F246A}$	$_{6His}Rbg1_{fl}$-$Tma46_{205-345\ I241A\ F246A}$*		
DRG_{fl}-$DFRP_{dfrp\ I241A}$	$_{6His}Rbg1_{fl}$-$Tma46_{205-345\ I241A}$*		
DRG_{fl}-$DFRP_{dfrp\ W249A\ K250E}$	$_{6His}Rbg1_{fl}$-$Tma46_{205-345\ W249A\ K250E}$*		
DRG_{fl}-$DFRP_{dfrp\ \Delta\alpha 1}$	$_{6His}Rbg1_{fl}$-$Tma46_{239-345}$*		

* Plasmids obtained from Dr. Bertrand Séraphin, IGBMC, France; [a] Plasmids obtained previously in the laboratory by Mercedes Spinola; [b] Plasmids obtained in the laboratory by Isabel Perez. Among the Rbg1-Tma46 mutants, $_{6His}Rbg1_{fl\ GFPSVGKS\ \rightarrow\ GFPSVGKN}$-$Tma46_{205-345}$ and $_{6His}Rbg1_{fl\ GFPSVGKS\rightarrow GFPSVAMN}$-$Tma46_{205-345}$ contain mutations in the G1 motif inactivating GTP binding/hydrolysis; Tma46, G272A R273E are mutations in the conserved residues with Gir2 and I241A, F246A, W249A and K250E are in the pi stacking interface with Rbg1 depicted in figure 4.12.

Cloning of the DRG_{s5d2l} constructs was done as described below. The plasmids $_{6His}Rbg1_{fl}$-$Tma46_{205-345}$, $_{6His}Rbg2_{fl}$-$Gir2_{174-238}$ and $_{6His}Drg1_{fl}$ were used to obtain $_{6His}Rbg1_{176-240}$, $_{6His}Rbg2_{174-240}$ and $_{6His}Drg1_{175-238}$ respectively. The insert DNA was amplified from the corresponding plasmids by polymerase chain reaction (PCR) using the Techne TC-3000

Thermal cycler. The reaction consisted of 2.5U of Pfu Turbo DNA polymerase (Agilent) in the commercial buffer (20 mM Tris-HCl, 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH4)_2SO_4$, 0.1% Triton X-100 and 0.1 mg/ml nuclease-free BSA), 0.2 mM dNTP mix, 0.4 µM oligonucleotide primer (synthesized by Sigma, given in table 3.2 with restriction enzyme sites) and 50-100 ng template DNA in 50 µl final volume. The PCR reaction was carried out as given: 94ºC for 30 sec followed by 30 cycles of amplification (Denaturation step of 94ºC for 30 sec and then 55ºC for 1 min of annealing followed by elongation step of 68ºC for 7 min) and finally at 68ºC for 10 min. The PCR product confirmed in 1% agarose gels was treated with DpnI for 1 hour at 37ºC for digestion of the template DNA and purified using PCR purification kit (GE Healthcare).

Cloning was done into pp-pET28a vector (Kanamycin resistance) which contains the cloned gene under expression control of a lac promoter and is a modification on pET28 with the replacement of the thrombin cleavage site for a PreScission protease site between the histidine tag and the cloned sequence (Obtained vector from Dr. Ramón Campos, (Campos-Olivas *et al.,* 2007)). Both insert and vector DNA were digested with the restriction enzymes NdeI and XhoI and purified prior to ligation in 25 µl using T4 DNA ligase (Roche) in the commercial buffer at 22ºC for 2 hours. Whole of the ligation reaction was transformed in 100 µl of chemically competent DH5α cells (Invitrogen) by the heat shock method.

Table 3.2: Primers used for cloning. The sequence of the oligonucleotides used for the cloning of DRG_{s5d2l} and $Rbg1_{fl\ GFPSV\mathbf{AMN}}$-$Tma46_{205-345}$ constructs. Also given are the restriction enzymes with their sites marked in italics in the primer sequence. The double mutation of $Rbg1_{fl\ GFPSV\mathbf{AMN}}$-$Tma46_{205-345}$ is shown as bold in the part where the forward and reverse primers overlap (given as capitals).

Construct	Restriction enzymes	Primer sequence
$Rbg1_{176-240}$	NdeI/XhoI	Fw: $^{5'}$CGTCTG*CATATG*ACTCCGCCAGATATCTTG$^{3'}$ Rv: $^{5'}$GGCAGG*CTCGAG*TCATCTTGACGAAGCTTCC$^{3'}$
$Rbg2_{174-240}$	NdeI/XhoI	Fw: $^{5'}$CGTCTG*CATATG*GAAAAACCAAACATTTATTAC$^{3'}$ Rv: $^{5'}$ACATTT*CTCGAG*TTAGCGATGTTGTTCGTTGATG$^{3'}$
$Drg1_{175-238}$	NdeI/XhoI	Fw: $^{5'}$CGCTTG*CATATG*AAACCCCCCAACATTGGC$^{3'}$ Rv: $^{5'}$ACAGGG*CTCGAG*TCATCTGTTTCCTTCCACCACATC$^{3'}$
$Rbg1_{fl\ GFPSV\mathbf{AMN}}$-$Tma46_{205-345}$		Fw: $^{5'}$GTCGGTG**GCGATG**AATACATTACtgtccaagttgactggtactgagtc$^{3'}$ Rv: $^{5'}$GTAATGTATT**CATCGC**CACCGACgggaacccgacaaaccccac$^{3'}$

Briefly, the DH5α-DNA mix was incubated in ice for 30 min and subjected to heat shock of 42ºC for 1 min 30 sec. Afterwards the reaction mixture was again kept in ice (5 min), added 200 µl of LB and incubated in a shaking incubator at 37ºC for 45-50 min. Finally, 150 µl of the cells were spread on LB agar plates containing 50 µg/ml kanamycin

and incubated overnight at 37ºC for the colonies to grow. Selected colonies were grown overnight in 5 ml LB media (with kanamycin) at 37ºC and plasmid DNA extracted using Qiaprep Spin Miniprep kit (Qiagen, following manufacturer´s instructions). The constructs were then confirmed by sequencing.

Mutagenesis of the G1 motif GFPSVGKS to GFPSV**AMN** was accomplished using a modified Quick Change Mutagenesis protocol. The $Rbg1_{fl\ S79N}$-$Tma46_{205-345}$ plasmid containing GFPSVGKS→GFPSVGK**N** mutation was used as the template for the PCR reaction. The oligonucleotides used for introducing the mutation are given in table 3.2. The PCR reaction was carried out essentially as described above except that *PfuUltra* II Fusion HS DNA Polymerase (Agilent) in *PfuUltra* II reaction buffer and only 10 ng of template DNA was used. The reaction was as follows: 95ºC 30 sec; 30 cycles of (95ºC 30 sec, 55ºC 30 sec, 68ºC 12 min); 68ºC 5 min. DpnI treatment was for 3 hours at 37ºC followed by 80ºC for 20 min to inactivate the enzyme. The product was then heat shock transformed in chemically competent MH1 cells. Plasmid DNA was purified using Miniprep kit (Macherey-Nagel) and the mutations confirmed by sequencing.

The rest of the plasmids, all of which contained kanamycin resistance unless otherwise mentioned, were heat shock transformed into DH5α (Invitrogen) chemically competent cells except that 50 ng of plasmid DNA (50 ng) was added to 60 µl of DH5α cells. The plasmids were subsequently purified using mini or midi prep kits (Qiagen) and stored at -20ºC. For expression, the plasmids were transformed in *E.coli* BL21-CodonPlus (DE3)-RIPL (Stratagene) chemically competent cells also using the heat shock method (42ºC for 25 sec) and plated in LB agar containing 50 µg/ml kanamycin and 33 µg/ml chloramphenicol. Glycerol stocks of the transformed bacteria were made with 20% glycerol, frozen with liquid nitrogen and stored in a -80ºC freezer.

3.1.2 Expression and purification of the different complexes

Adequate expression of the proteins was first assured by small scale expression in 100 ml cultures (LB media containing 50 µg/ml kanamycin and 33 µg/ml chloramphenicol) induced with 1 mM IPTG, followed by batch purification using affinity Ni-NTA resin and confirmed by gel electrophoresis.

Large scale expression of the recombinant proteins, if not mentioned otherwise, was undertaken in autoinduction media following the Studier (Studier 2005) protocol. The composition of the media was as given in table 3.3. A pre-culture was grown in 5ml of ZYP-0.8G containing appropriate antibiotic starting from the glycerol stock at 37ºC overnight in a shaking incubator (Innova 43,New Brunswick scientific) at 250 rpm. The next day the preculture was inoculated into 500 ml of ZYP-5052 containing appropriate antibiotic in 2 litre conical flask and grown for 6-7 hours. When the OD_{600nm} reached

around 0.8-1.0, the temperature was reduced to 20ºC and left overnight for expression of the protein. Afterwards, the cells were harvested by centrifugation at 4,000 rpm (4547 g) for 20 min in a Beckman Coulter J6-HC centrifuge using a JS-4.2A rotor. The pellet was drained of the media and resuspended in 40 ml of PBS 1x and centrifuged again in an Eppendorf Centrifuge 5810R at 4000 rpm (3220 g) for 20 min at 4ºC. The supernatant was discarded and the pellets were then frozen with liquid nitrogen and stored in a -80ºC freezer until further use.

Table 3.3: Autoinduction media components.

Stock Solutions	ZYP-0.8G Final concentration	ZYP-5052 Final concentration
ZY	To make up volume	To make up volume
1M $MgSO_4$	1mM	1mM
10000x metals mix*	1x	1x
40% glucose	0.5%	-
50x 5052*	-	1x
20x NPS*	1x	1x
Kanamycin	50 µg/ml	50 µg/ml
Chloramphenicol	33 µg/ml	33 µg/ml

* Recipes and full forms of the abbreviations are given in the appendix section 8.1 and 8.2 respectively.

For the purification of the different proteins, the *E. coli* cell pellet was thawed in ice and mixed with lysis buffer (300 mM NaCl, 20 mM Imidazole, 1 mM β-Mercaptoethanol (β-MeOH), 50 mM Tris pH 8.0 and 1 tablet of protease inhibitor cocktail (Complete, EDTA-free, Roche) and made up to a total volume of 50 ml. The cells were lysed by sonication (Bioblock Scientific Vibracell 75042) at 30% amplitude for 10 min with pulses of 1 sec ON and 1 sec OFF. The solution for lysis was kept in ice during the sonication to prevent overheating and damage to the protein. The lysate was cleared by centrifugation at 16,000 rpm (30651 g) (Sorvall RC-6 centrifuge using an SS-34 rotor) at 4ºC for 30 min and filtering through a 0.45 µm sterile membrane filter.

The filtered solution was then loaded into an affinity HisTrap FF column (GE Healthcare). The column contained 5 ml bed volume of nickel-nitriloacetic acid (Ni-NTA) resin (the nickel being reloaded by passing 1 volume of 100 mM $NiSO_4$ solution in case of regenerated columns). The protein was loaded onto this column, previously equilibrated with Buffer A (300 mM NaCl, 20 mM Imidazole, 1 mM β-MeOH, 50mM Tris pH 8.0), using a MiniPuls 3 peristaltic pump (Gilson) at 4ºC at a flow rate of 1-3 ml/min. After washing the column with 10 bed volumes of Buffer A, the protein in the column was eluted out with a gradient of imidazole (20-500 mM Imidazole) using an ÄKTA Purifier (GE Healthcare). The different fractions from the nickel column were analysed by SDS polyacrylamide gel electrophoresis and the ones containing the protein were retrieved and kept in ice. These fractions were pooled and concentrated to a final volume of 1-2 ml by

centrifugation at 4000 rpm (3220 g) at 4ºC, using Amicon Ultra centrifugal filter devices of adequate pore size (generally 3000, 10000 or 30000 molecular weight cut-offs).

The concentrate was directly loaded onto a pre-equilibrated size exclusion column (Sephadex 200 or 75 (16/60 or 26/60) columns) and the protein eluted at around 1.0 ml/min using an ÄKTA Prime system (GE Healthcare) located in a 4ºC cabinet. The size exclusion buffer (Buffer S) with the composition 150 mM NaCl, 20 mM Tris pH 7.5 and 2 mM DTT was used. The eluted protein fractions from this final step of purification were then pooled together and concentrated by ultrafiltration using the Amicon concentrator at 4ºC, flash-frozen with liquid nitrogen and stored at -80ºC until further use.

3.1.3 Protein detection and quantification

Detection of the proteins at each step of the expression and purification procedure was done using Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE). The samples for the gel were prepared by adding 1x Protein Loading Buffer (1% w/v SDS, 62.5 mM Tris pH 6.8, 8% v/v Glycerol, 0.03% v/v Bromophenol Blue, 9% β-MeOH) and boiling at 99ºC for 6-10 min. Then 10-20 µl of each sample along with the standard were run in 10% or 12% polyacrylamide gels with 1x MES-SDS running buffer. For visualization, the gels were stained with Coomassie brilliant blue-R250 for 10-30 min in staining solution (0.1% w/v Coomassie blue in 10% v/v acetic acid and 40% v/v ethanol) followed by de-staining in a solution of 10% v/v acetic acid and 40% v/v ethanol.

The concentrations of the proteins were determined by the Bradford method (Bradford 1976) using the Bio-Rad protein assay solution against Bovine Serum Albumin (BSA) standard.

3.1.4 Native gel electrophoresis for detection of protein-heavy metal binding

For the native gel band shift assay, the protein at 3 mg/ml (~50 µM) concentration was first incubated with 5-10 mM of heavy metal solutions prepared in milliQ water for about 30 min in ice. To these samples, 4% glycerol was added and loaded in a 7% native polyacrylamide gel using glycine running buffer (50 mM glycine adjusted to pH 9 using NaOH). The voltage used was 150V during 3-5 hours at 4ºC. Staining and destaining were done as described in section 3.1.3.

3.1.5 Production of Selenomethionine derivative of $Rbg1_{fl}$–$Tma46_{205-345}$

Selenomethionine-substituted protein crystals were obtained by incorporating selenomethionine into the $Rbg1_{fl}$–$Tma46_{205-345}$ protein using the autoinduction method described by Sylvie Doublie (Doublie 2007, Sreenath *et al.,* 2005).

Briefly a single colony of BL21-CodonPlus(DE3)-RIPL with plasmid $Rbg1_{fl}$–$Tma46_{205-345}$ was inoculated into 3 ml of PA-0.5G containing antibiotics (50 µg/ml Kanamycin and 33 µg/ml chloramphenicol) and grown at 37ºC at 250 rpm for 7-8 hours in a shaking incubator (Innova 43, New Brunswick scientific). The temperature was decreased to 25ºC and the preculture left for growing for a further 30 min. This was then inoculated onto 100 ml PA-0.5G containing antibiotics and grown for 18 h at 25ºC, 250 rpm. 20 ml of this culture was inoculated to 500 ml PASM-5052 in 2 litre flasks and grown at 25ºC, 250 rpm. After 24 hours of growth, the culture was centrifuged and pelleted as before and frozen in liquid nitrogen and stored in a -80ºC freezer until further use.

Table 3.4: Autoinduction media components for the expression of the selenomethionine-labelled protein.

Stock Solutions	**PA-0.5G** **Final concentration**	**PASM-5052** **Final concentration**
Deionised H_2O	To make up volume	To make up volume
1M $MgSO_4$	1mM	1mM
10000x metals mix*	1x	1x
40% glucose	0.5%	-
50x 5052*	-	1x
20x NPS*	1x	1x
50x Aminoacids mixture*	1x	1x
L-methionine solution (25 mg/ml)	0.1 mg/ml	0.01 mg/ml
L-Selenomethionine solution (25 mg/ml)	-	0.125 mg/ml
1000x Vitamins solution*	1x (with Vitamin B_{12})	1x (lacking Vitamin B_{12})
Kanamycin	50 µg/ml	50 µg/ml
Chloramphenicol	33 µg/ml	33 µg/ml

* Recipes and full forms of the abbreviations are given in the appendix section 8.1 and 8.2 respectively.

The purification protocol for the selenomethionine-substituted $Rbg1_{fl}$–$Tma46_{205-345}$ protein was the same as that for the native protein except in the size exclusion buffer, the concentration of DTT was increased to 5 mM in the size exclusion buffer to prevent oxidation of selenomethionine.

3.2 Crystallization of $Rbg1_{fl}$–$Tma46_{205-345}$

3.2.1 Crystallization trials

The purified protein samples ranging in concentration from 20-40 mg/ml were used for setting up the crystallization drops. The initial crystallization trials were set up with the sitting drop method in 96-well plates using the commercial crystal screens available. These included Hampton (Crystal Screen, Index), Jena Bioscience (HTS I & HTS II), Qiagen

(Nextal Classics & PEGS, JCSG+, PACT) and Emerald BioSystems (Wizard I & II) among others. The screening studies explored varying parameters comprising of protein concentration, nature and concentration of precipitant, pH, temperature etc. Plates were set up in a high-throughput manner in 96 well MRC crystallization plates (Molecular Dimensions MD11-00-100) using multi-channel pipettes or with robotic workstations (Honeybee Cartesian MycroSys synQUAD liquid handling robot from Genomic Solutions). The drop solution contained the protein/reservoir solution usually in 1:1 ratio with a final drop size of 0.6 µl. For each condition, the plates were incubated inside a 21ºC temperature controlled room. Further optimization trials were performed in 24 well plates (NeXtal-Qiagen) and also at 4ºC. The volume of reservoir solution used here was 500 µl with a drop size of 2 µl at a protein/reservoir ratio of 1:1.

All crystals of $Rbg1_{fl}$-$Tma46_{205-345}$ complex were cryoprotected by soaking in a solution of the reservoir made up with 1.9 M Sodium malonate pH 7.0 (CP-Buffer-8, AxyGen Biosciences) and flash-cooled by either introducing swiftly into a cryostream of nitrogen gas at 110K or by plunging into liquid nitrogen.

3.2.2 Crystal data collection, structure resolution, refinement and analysis

The X-ray diffraction data were collected from single flash cooled crystals at the beam line ID14-4 (McCarthy *et al.,* 2009) at the European Synchrotron Radiation Facility (ESRF) at Grenoble, France. The diffraction images were obtained with ADSC Quantum Q315r CCD detector.

With the purpose of improving the diffraction quality of the crystals obtained, crystal dehydration experiments were also carried out. This was done using the Humidity Control Device (HC1) at the BM-14 beamline in the European Synchrotron Radiation Facility (ESRF), Grenoble which can dehydrate crystals in a controlled manner. The crystals from the reservoir solution were directly mounted in Mitegen Micromesh 400/10 µm or 400/25 µm loops at room temperature without cryoprotection. The crystals were then introduced into the humidified air stream and the relative humidity was varied in steps of 1%, collecting diffraction images at each step.

All images collected were processed using the programs available in the CCP4 suite (Collaborative Computational Project, Number 4, (CCP4 1994)). iMOSFLM was used to index and integrate the images and the intensities were then merged and scaled using SCALA module in CCP4. Pointless and Self Rotation Function from Molrep were used to further estimate the symmetry of the crystal. Heavy atom site searching and phasing were done by autoSHARP (Fortelle and Bricogne 1997) or SHELX implemented in HKL2MAP (Pape and Schneider 2004, Sheldrick 2008). The initial model building/tracing was done with ARP/wARP (Langer *et al.,* 2008) as implemented in autoSHARP. Cycles of manual

model building were performed with the program Coot (Emsley and Cowtan 2004). Waters belonging to the first solvation shell were added to the model using ARP/wARP (CCP4) and validated with the electron density maps in Coot. The structures were refined with REFMAC (Murshudov *et al.,* 1997) for isotropic refinement. TLS groups were defined and used for anisotropic refinement. Superpositions between the structures were done using the SSM superpose function (Krissinel and Henrick 2004) in Coot and analysis of the electrostatic surface potential was performed using APBS (Baker *et al.,* 2001) in Pymol, also used for generating the structure figures (Schrodinger 2010). Structural similarity searches were performed with the DALI webserver (Holm and Rosenstrom 2010).

3.2.3 Crystallization and structure determination of $Rbg1_{fl\ GFPSVGKN}$-$Tma46_{205-345}$ and $Rbg1_{fl}$-$Tma46_{205-345}$ - GDP/GTPγS complex

Crystals of the inactive GTPase mutant ($Rbg1_{fl\ GFPSVGKN}$-$Tma46_{205-345}$) and the wild type protein complex bound to GDP/GTPγS were obtained in the same conditions as of the native $Rbg1_{fl}$ -$Tma46_{205-345}$ complex, i.e 2.38 M Sodium formate, 0.2-0.25 M Tri-sodium citrate dehydrate pH 6.5 incubated at 4ºC in 24-well NeXtal plates. 1-5 mM of GDP/GTPS was used for the co-crystallization of the protein-nucleotide complex. Around 60 mg/ml of protein concentration was used in a 1:1 ratio of protein to reservoir solution. Flash frozen crystals cryoprotected with 1.9 M Sodium malonate were diffracted at I03 at the Diamond Light Source, UK.

Briefly, the diffraction images were processed using iMOSFLM followed by SCALA. Rigid body refinement was performed in REFMAC. TLS refinements followed along with visualization in Coot as used for the native $Rbg1_{fl}$-$Tma46_{205-345}$ structure (Refer section 3.2.2 and 4.3).

3.3 GTP binding and hydrolysis assays

3.3.1 Differential scanning fluorimetry for protein stability and GTP binding

Fluorescence based thermal shift assay was performed using protein samples at around 0.05 mM in buffer S (Size exclusion buffer) with 5x Sypro Orange, with or without 0.02 mM GDP, GTP or 0.02/0.05 mM GTPγS in wells of MicroAmp 96-well Fast Optical Reaction plate (Applied Biosystems). The plates were sealed with MicroAmp Optical sealing tape (Applied Biosystems), given a spin down at 1000 rpm for 1 min and heated from 20 to 85ºC in increments of 1ºC in a T7500 Fast Real-Time PCR System (Applied Biosystems). The fluorescence readings were taken in duplicate. The wavelengths for excitation and emission were 490 nm and 550 nm respectively. The results were analysed and visualized graphically in GraphPad Prism 4 software (GraphPad Software Inc.).

3.3.2 Malachite green assay for GTP hydrolysis

Numerous highly sensitive, non-radioactive assays for phosphate measurement have relied upon Malachite Green reagent and are based on the formation of a colorimetric complex between the basic dye malachite green and phosphomolybdate (Baykov *et al.,* 1988). Briefly the malachite green complexes with acidified ammonium molybdate and in the presence of inorganic phosphate form a green complex which has an absorbance at 630 nm range which can be measured using a UV-Visible spectrophotometer. The reagent components Malachite Green carbinol hydrochloride, Ammonium molybdate, Polyvinyl alcohol, GTP and GDP were obtained from Sigma Aldrich (213020, A7302, P8136, G9002, G7127 respectively). GTP and GDP 100 mM stock solutions were made in Milli-Q water (Millipore corporation) and stored at -20ºC for further use. The assay reagent components: 5.72% w/v ammonium molybdate in 6N HCl, 0.08% w/v malachite green solution, 2.32% w/v polyvinyl alcohol are prepared individually and stored at 4ºC. On the day of the assay, the components are mixed in a ratio of 2:2:1:1 for MilliQ water: Malachite green: Polyvinyl alcohol: Ammonium molybdate and the complex allowed to mature to a yellow colour for 3 hours.

The protein samples were prepared at a concentration of 20 µM in filtered and degassed buffer containing 100 mM KCl, 50 mM Tris pH 7.5 and 5 mM $MgCl_2$. Reaction was done in 50 µl volumes in Microtest 96-Well Flat Bottom plates (Sarstedt). A phosphate standard was prepared from KH_2PO_4 to quantify the inorganic phosphate released from the samples. Blank with no protein in the reaction mixture was subtracted from the protein sample readings. After incubating for 1 h at 37ºC, 200 µl of Malachite green reagent was immediately added onto all the wells using a multipipette and after 5 min of colour development, the absorbance readings were measured in a Wallac Victor2 1420 Multilabel Counter using a 630 nm filter. The data were fitted to Michaelis-Menten equation using non-linear regression in GraphPad Prism 4 (GraphPad Software Inc.) to determine the kinetics. The Mg^{2+} dependence assay was carried out with 0, 5 and 10 mM $MgCl_2$ in the assay buffer.

3.4 RNA homopolymer binding assay

Nonspecific RNA binding studies were carried out by incubating each protein with polyuridylic acid-Agarose (matrix - polyacrylhydrazido-agarose) beads (Sigma Ref. No. P8563). Around 10 µl of 50% poly(U)-agarose beads with 0.2-0.4 mg of protein were nutated in a 4ºC room for 30 min in 50 µl of reaction buffer containing 100mM NaCl, 10 mM Hepes pH 7.4, 2 mM $MgCl_2$, 0.1% Triton X-100, 3 mM DTT and 0.1 mg/ml BSA. After incubation the beads were pelleted down at 10,000 rpm (g) for 1 min in a Sorvall Legend 21R microcentrifuge, the supernatant discarded and the pellet resuspended in 500 µl of wash buffer (reaction buffer without BSA). The beads were washed six times in this

manner and finally the poly(U)-agarose bound proteins were eluted by adding 10 µl of SDS sample buffer and boiled at 95ºC for 8 min. The samples were then run on a 10% SDS-polyacrylamide gel and visualized by Coomassie blue staining. BSA and poly(U)-agarose beads treated with 1 mg/ml of RNase A added prior to the addition of the protein were used as negative control. Binding in presence of free polyuridylic acid (Sigma Ref. No. P9528) was performed by incubating free poly(U) at concentrations of 0, 0.1 and 1 mg/ml with the protein in the reaction buffer for 20 min at 4ºC prior to the addition of the beads and elution. Heparin at concentrations of 0.5, 1, 2 and 4 mg/ml were used in the reaction mixture prior to the addition of the beads to analyse the strength of binding.

3.5 Yeast experimental study

3.5.1 Yeast strains used

Saccharomyces cerevisiae strains used were derived from BMA64 (Baudin-Baillieu *et al.,* 1997). The gene deletions introduced are as given in table 3.5.

Table 3.5: Genotype of the wild type and mutant yeast strains.

Name	Strain	Relevant genotype
BSY1664*	Wild type (reference)	*MATa ade 2-1 his3-11,15 leu2-3,112 Δtrp1 ura3-1 can1-100*
BSY1174*	Mating type	Ade 5,rho+ mit- (oxi3) 43 : M8-227-1/1
BSY1177*	Mating type	Ilv5, trp2 (leaky), omega-, PR 61 :M12-54
BSY2049*	Δ*rbg1*Δ*rbg2*Δ*slh1*	*MATα Δrbg1::HISMX6 Δrbg2::KanMX4 Δslh1::URA3Kl*
BSY2057*	Δ*tma46*Δ*gir2*Δ*slh1*	*MATa Δtma46::HISMX6 Δgir2::KanMX4 Δslh1::TRPKl*
BSY2077*	Δ*rbg2*Δ*slh1*	*MATa Δrbg2::KanMX4 Δslh1::URA3Kl*
BSY2095*	Δ*rbg1*Δ*tma46*Δ*slh1*	*MATα Δrbg1::HISMX6 Δtma46::HISMX6 Δslh1::URA3Kl*
BSY2671	Δ*rbg1*Δ*rbg2*Δ*tma46*Δ*slh1*	*MATa Δrbg1::HISMX6 Δrbg2::KanMX4 Δtma46::HISMX6 Δslh1::URA3Kl*
BSY2672	Δ*rbg1*Δ*rbg2*Δ*tma46*Δ*slh1*	*MATα Δrbg1::HISMX6 Δrbg2::KanMX4 Δtma46::HISMX6 Δslh1::URA3Kl*

* Strains obtained from Dr. Bertrand Séraphin, IGBMC.

All strains were grown in standard YPD (yeast extract/peptone/dextrose) liquid media at 30°C in Infors HT Multitron Standard shakers at 170 rpm or on YPDA or synthetic medium plates.

3.5.2 Crossing of yeast strains

The quadruple gene deletion mutant Δ*rbg1*Δ*rbg2*Δ*tma46*Δ*slh1* was constructed by crossing the two haploid parental yeast strains containing respectively the triple gene deletion, Δ*rbg1*Δ*tma46*Δ*slh1* and the double-gene deletion, Δ*rbg2*Δ*slh1* obtained from Bertrand Séraphin (IGBMC). For this, both the strains were patched together in a YPDA (yeast extract/peptone/dextrose/agar) plate and left overnight in a 30ºC incubator allowing the diploids to form. The next day, the plate was replica-plated onto selective complete medium (CSM) lacking histidine and kept overnight at 30ºC. This was then replica-plated onto YPDA plates containing G418 (Geneticin antibiotic, 200 µg/ml) and incubated

overnight at 30ºC. The surviving diploids from the YPDA-G418 plate were replica-plated onto a sporulation plate containing aminoacids (Sporulation media contains 1% potassium acetate pH 7.0, 0.1% Bacto-yeast extract, 0.05% glucose, 2% Bacto-agar). The sporulation plate was incubated at 25ºC for one week until sufficient number of tetrads was seen as observed under a microscope.

3.5.3 Tetrad dissection

Tetrads were dissected on a Singer MSM-400 Dissection microscope. For this, a small amount of Zymolyase 100T (Seikagaku, Japan) was dissolved in 1 ml of sterile SCE buffer (1M Sorbitol, 0.1M Sodium citrate pH 7.0, 60 mM EDTA). A small colony size amount of sporulated yeast was taken and resuspended in 50 µl of SCE/Zymolyase mix for 7-10 min. The reaction was stopped by diluting to 500 µl with sterile water. 30 µl of the diluted cells were spread on one side of a YPDA plate and dissected using a micromanipulator (Singer MSM) under the microscope. The resulting plates were incubated at 25°C until the haploid spores gave colonies of a reasonable size which were streaked onto a master plate (YPDA). The master plate was then replica-plated to various selective media: Complete Synthetic Media CSM (-Leu/-Ade/-His/-Trp/-Ura), YPDA with G418 (200 µg/ml), synthetic minimal selective media plated (0.67% Yeast Nitrogen Base without aminoacids and without ammonium sulfate, 2% Glucose, 2% Bacto-agar) with the mating type tester strains (MATa or MATα). The replica plates were scored after one day of incubation at 25°C and the haploid corresponding to the quadruple mutant (both MATa and MATα) were identified. Glycerol stocks were made from 2 ml YPD overnight cultures (30ºC) by mixing 300 µl of sterile glycerol (100%) with 700 µl of the culture and stored at -80ºC.

3.5.4 Transformation of plasmids into yeast

Starting from colonies, a 50 ml yeast culture (mutants and wild type) was grown to an OD_{600} of 0.5-0.8, centrifuged and the yeast cells washed twice, first with 50 ml sterile 10 mM Tris-HCl pH 7.5 and then with 25 ml of sterile LiT buffer (10 mM Tris-HCl pH 7.5, 100 mM lithium acetate). The final pellet was resuspended in sterile LiT to a final volume of 800 µl also adding sterile DTT to 10 mM. 100 µl of the competent yeast cells were added to carrier DNA/LiT mix (50 µl of LiT, 5 µl of salmon sperm carrier DNA 10 mg/ml) and the plasmid DNA (0.5-2 µg) and incubated at room temperature for 10 min. 300 µl of a PEG/LiT mix (sterile 50% PEG4000 in LiT) was then added and kept at 30ºC for another 10 min. This was followed by incubation at 42°C for 15 min after which the cells were centrifuged for 1 min at around 2000 g. The supernatant was discarded and 1 ml of YPD was added and incubated at 30 °C for 60 min at 1400 rpm in a Thermomixer (Eppendorf). The cells were spun down and resuspended in 200 µl of 10 mM Tris-HCl pH 7.5. In

appropriate selective plates (CSM-Leu, CSM-Ura or CSM-Leu-Ura), 50 µl was plated and incubated at 30°C until colonies appeared.

3.5.5 Yeast growth phenotype assay

Comparative growth rate of the wild type and mutant strains were assayed using the drop assay. Briefly, yeast cultures were grown overnight to an OD_{600} of 2-3. Dilutions of 10-fold were made starting from an optical cell density (OD_{600}) of 0.1. On CSM-Leu plates, 3 µl of each dilution was spotted in an identical manner and incubated at 30° and 37°C. The plates were photographed once sufficient growth was observed and the number of days of incubation at each temperature recorded.

3.5.6 Translation reporter assays

Gene reporter assays have been widely used to measure transcriptional or translational activity of cells *in vivo*. The translation reporter assays done here were performed using the Dual-Luciferase® Reporter Gene assay system (DLR^{TM}, Promega) which measures the bioluminescence due to firefly (*Photinus pyralis*) and *Renilla* (*Renilla reniformis*/sea pansy*)* luciferase enzyme activities. Luciferases are oxidative enzymes which during a chemical reaction generate light through the release of chemical energy in their substrates as photons. Accordingly, firefly luciferase catalyses the formation of light from ATP and luciferin as per the following reaction:

$$\textit{Luciferin} + \textit{ATP} + O_2 \xrightarrow[Mg^{2+}]{\text{Firefly luciferase}} \text{Oxyluciferin} + \text{AMP} + \text{PPi} + CO_2 + \text{Light}$$

and *Renilla* luciferase catalyses the conversion of coelenterazine:

$$\text{Coelenterazine} + O_2 \xrightarrow{\textit{Renilla}\ \text{luciferase}} \text{Coelenteramide} + CO_2 + \text{Light}$$

The process of luminescence is different from fluorescence in that no excitation is needed and involves a chemical reaction. The luciferase enzymes have emission spectra with broad peaks. The peak emission spectrum for the firefly luciferase can vary between 560 nm (yellow-green) and 620 nm (red) and that for *Renilla* is about 480 nm (blue light). Hence, the measurement is done over a wide range of visible spectrum. The bioluminescence emitted from the firefly luciferase (Fluc) and *Renilla* luciferase (Rluc) were measured using a Lumat LB 9507 luminometer (Berthold Technologies) which contains a photomultiplier tube operated in single photon counting mode (380-630 nm).

Table 3.6: The dual luciferase reporter plasmids used for the stop codon recognition and translational fidelity assays. The first two plasmids were used for the stop codon recognition assay, one of them containing a stop codon (UAA) between the *Renilla* and firefly luciferase genes and the other containing a glutamine (CAA) instead of the stop codon. The rest of the plasmids were used for the translation fidelity assay with mutations replacing K529 of the firefly gene (Kramer and Farabaugh 2007, Kramer *et al.,* 2010).

Name	Mutated to codon	Codes for	Position
pBS3564	CAA	Glu	Stop codon recognition assay. Mutation in between *Renilla* and firefly genes
pBS3565	UAA	Stop	
pBS4082	UUU	Phe	Translational fidelity assay. Mutation at lysine-529 of firefly gene
pBS4083	UAA	Stop	
pBS4084	UAG		
pBS4085	UGA		
pBS4086	CAA	Gln	
pBS4087	CAG		
pBS4088	GAG	Glu	
pBS4089	AUA	Ile	
pBS4090	AUG	Met	
pBS4091	ACA	Thr	
pBS4092	ACG		
pBS4093	AGA	Arg	
pBS4094	AGG		
pBS4095	AAU	Asn	
pBS4096	AAC		
pBS4097	AAA	Lys	

For the assay, the Δ*rbg1*Δ*tma46*Δ*slh1* mutant and wildtype (BSY1664) yeast strains were transformed with the indicated reporter plasmids given in table 3.6. These yeast expression vectors contained the *Renilla* and firefly luciferase genes fused into a single open reading frame (Rluc-Fluc) under the transcriptional control of a *PGK* promoter and encoded a single bifunctional protein. The yeast transformation protocol was as described in section 3.5.4. The reporter plasmids contained the Ura marker and hence the colonies were selected on CSM-Ura plates. In the case of transformations with the Δ*rbg1*Δ*tma46*Δ*slh1* strain complemented with the wild type Rbg1 (HA-tagged) protein (Leu marker), colonies were selected on CSM-Ura-Leu plates. Starting from the colonies of the transformed wild type and mutant strains, 2 ml cultures were grown to a cell density (OD_{600}) of 0.5-0.8 at 30ºC. The extract preparation and the dual luciferase assay were performed essentially following the manufacturer's instructions (Promega Dual Luciferase assay manual).

Briefly, for each cell extract, the dual luciferase assay was performed by sequentially measuring the firefly and *Renilla* luciferase activities from 5 µl of the same extract in the luminometer. First, the firefly luciferase activity was measured by adding the Stop and Glo reagent containing luciferin (Promega manual) which generates a glow-type

luminescent signal. The luciferase reaction is then quenched and the Lar II reagent for the *Renilla* reaction is added and the bioluminescence measured. Two such measurements were done for each cell extract and the values expressed in relative light units (RLU) were averaged. Assay for the stop codon recognition was done in triplicate, i.e., from three different cultures starting from the colonies while that for the fidelity assay was measured in duplicate; the result being given as the mean.

The emitted light intensity is directly proportional to the amount of luciferase enzyme present and the luminescence from *Renilla* is used as an internal control for normalization of gene expression. Thus the standardized Fluc activity was expressed as the ratio of firefly to *Renilla* luciferase activity (Fluc/Rluc). For the assay measuring the translational misreading error by the $tRNA_{UUU}^{Lys}$ for the lysine 529 of firefly luciferase, the standardized Fluc activity of the samples containing the mutant codon at position 529 of firefly were calculated relative to the sample containing the wild type lysine at this position. For the stop codon recognition assay, the level of read-through (stop codon suppression) associated with the mutation to glutamine (CAA) in place of the stop codon (UAA) in between the *Renilla* and firefly enzymes was calculated using the following equation:

$$Percentage\ readthrough = \frac{\frac{Fluc\ UAA}{Rluc\ UAA}}{\frac{Fluc\ CAA}{Rluc\ CAA}} \times 100$$

Results

4. Results

4.1 DRG-DFRP protein expression and purification

In order to better understand the structurally and functionally elusive DRG-DFRP protein complexes, the yeast Drg1-Dfrp1/Drg2-Dfrp2 proteins (Rbg1-Tma46, Rbg2-Gir2 respectively) and the human Drg1-Dfrp1 (Drg1-Lerepo4) proteins were expressed and purified for subsequent crystallization trials and GTP/RNA binding assays. The various constructs are given in table 3.1. All constructs of Rbg1 with Tma46 and Rbg2 with Gir2 were obtained from Dr. Bertrand Séraphin as part of the 3D Repertoire project. The constructs of Drg1 and Lerepo4 were made respectively by Mercedes Spinola and Isabel Perez in the laboratory.

The different proteins were essentially over expressed in *E.coli* expression system using the autoinduction method unless otherwise specified (Refer section 3.1.2, table 3.3). All individual proteins had a 6-histidine tag at the N-terminal except when expressed along with its binding partner in which case only one of them carried the His tag (as given in table 3.1). The individual or complex proteins were isolated from the cell lysate by chelating affinity column chromatography loaded with nickel, by virtue of the presence of the 6-His tag which binds to the nickel ions in the column resin. Upon applying a gradient of 20 to 500 mM Imidazole which competes with the histidines for nickel binding sites, the proteins generally eluted out at around 50-100 mM of Imidazole concentration in the buffer.

All of the proteins could be efficiently extracted from the bacterial lysate in this manner demonstrating that the His tags at the N-terminal were accessible. The protein containing fractions from the nickel affinity chromatography as visualized in a polyacrylamide gel (SDS-PAGE) were concentrated and purified using a size exclusion column to achieve structurally homogenous, crystallization-grade, pure protein.

The DRG-DFRP protein complexes always eluted as a single heterodimer in size exclusion chromatography. Apart from being a strong binding partner, DFRP proteins have been thought to be necessary for the functional and physical existence of the DRG proteins (Ishikawa *et al.,* 2005). Consistent with this observation, although $Rbg1_{fl}$ and $Drg1_{fl}$ could be overexpressed individually in bacteria and purified as shown in lanes 4 and 7 respectively in figure 4.1.C, $Rbg2_{fl}$ could not be expressed alone. However, when cotransformed with $Gir2_{fl}$, the $Rbg2_{fl}$-$Gir2_{fl}$ complex could be obtained which suggests that Rbg2 could not have been expressed or degraded upon expression in the absence of $Gir2_{fl}$. Full length Tma46 on the other hand, could not be overexpressed in *E. coli* and degrades when expressed even as a complex with Rbg1. However, several stable C-terminal truncations of Tma46 in complex with Rbg1 could be expressed and purified as given for example in lanes 1-3 in figure 4.1.C.

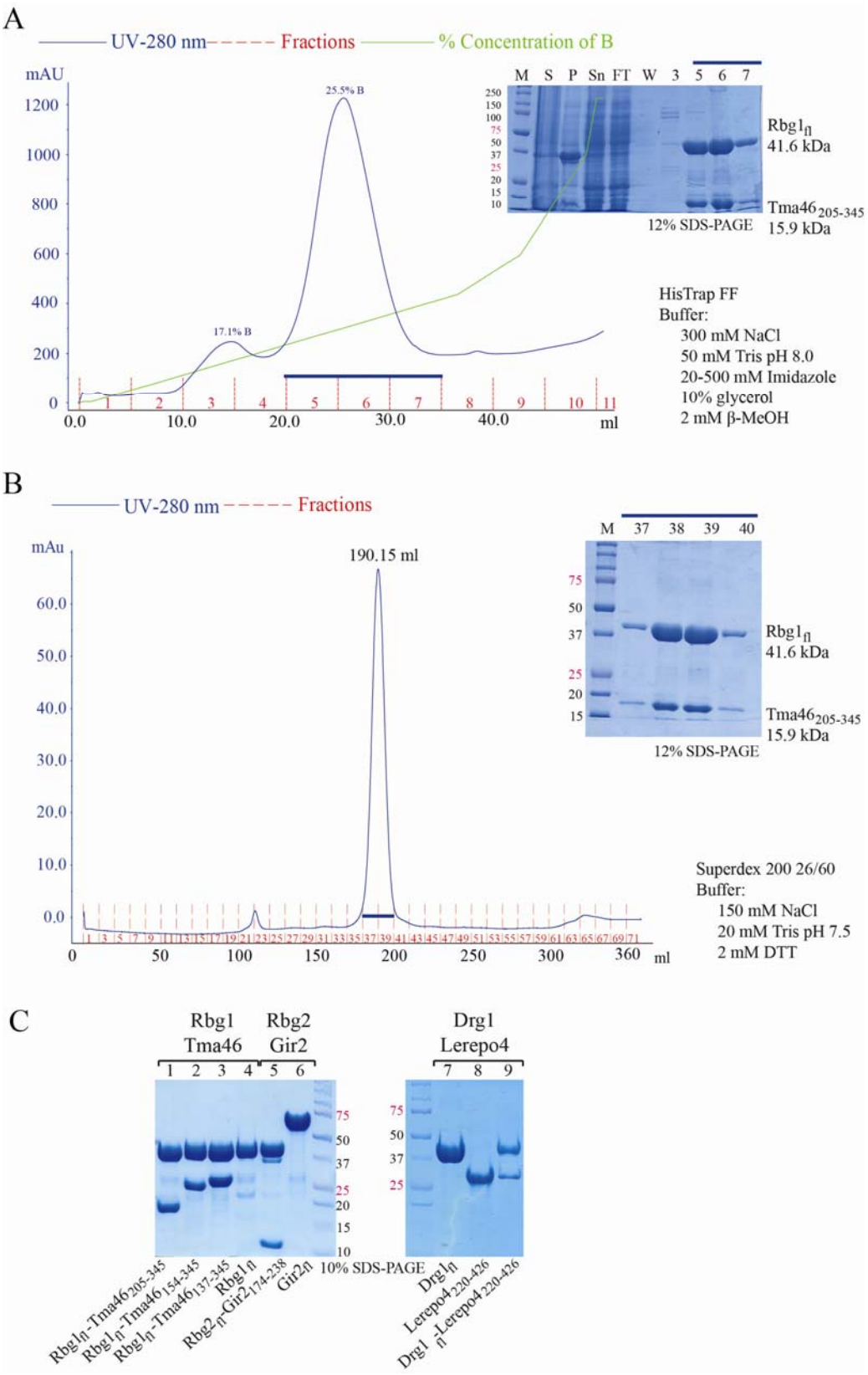

A
UV-280 nm
Fractions
% Concentration of B
mAU
25.5% B
17.1% B
M S P Sn FT W 3 5 6 7
Rbg1fl
41.6 kDa
Tma46205-345
15.9 kDa
12% SDS-PAGE
HisTrap FF
Buffer:
300 mM NaCl
50 mM Tris pH 8.0
20-500 mM Imidazole
10% glycerol
2 mM β-MeOH
ml
B
UV-280 nm
Fractions
mAu
190.15 ml
M 37 38 39 40
Rbg1fl
41.6 kDa
Tma46205-345
15.9 kDa
12% SDS-PAGE
Superdex 200 26/60
Buffer:
150 mM NaCl
20 mM Tris pH 7.5
2 mM DTT
ml
C
Rbg1
Tma46
Rbg2
Gir2
Drg1
Lerepo4
10% SDS-PAGE
Rbg1fl-Tma46205-345
Rbg1fl-Tma46154-345
Rbg1fl-Tma46137-345
Rbg1fl
Rbg2fl-Gir2174-238
Gir2fl
Drg1fl
Lerepo4220-426
Drg1fl-Lerepo4220-426

Figure 4.1: Purification results for the DRG-DFRP complexes.

A. Chromatogram for the first purification step of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex using the nickel chelating HisTrap fast flow column. The peaks correspond to the elution of the protein given by the UV absorption at 280 nm (blue line). The increasing percentage concentration of buffer B (containing 500 mM Imidazole) during the gradient is given as the green line. The fractions 5, 6 and 7 corresponding to the peak (25.5% B) was found to contain the $Rbg1_{fl}$-$Tma46_{205-345}$ complex as visualized in the Coomassie-stained SDS-PAGE gel and was concentrated.

B. The concentrated $Rbg1_{fl}$-$Tma46_{205-345}$ protein sample was then loaded in a Superdex 200 26/60 column. The size exclusion chromatogram shows the elution of the protein at 190.15 ml which corresponds to a molecular weight of ~70 kDa (MW of complex is 57.5 kDa and so suggests a single heterodimer). The fractions 37-40 were concentrated and stored at -80ºC.

C. The concentrated samples after purification by affinity followed by size exclusion chromatography techniques of the proteins $Rbg1_{fl}$-$Tma46_{205-345}$ (as representative), $Rbg1_{fl}$-$Tma46_{154-345}$, $Rbg1_{fl}$-$Tma46_{137-345}$, $Rbg1_{fl}$, $Rbg2_{fl}$-$Gir2_{174-238}$, $Gir2_{fl}$, $Drg1_{fl}$, $Lerepo4_{220-426}$ and $Drg1_{fl}$-$Lerepo4_{220-426}$ is given. The latter $Drg1_{fl}$-$Lerepo4_{220-426}$ complex was obtained by mixing together the individually expressed and affinity purified $Drg1_{fl}$ and $Lerepo4_{220-426}$ proteins and loading in the size exclusion column and separately concentrating the fractions containing the complex. The marker used for the protein gel was Precision plus Protein Dual Color Standards (BioRad).

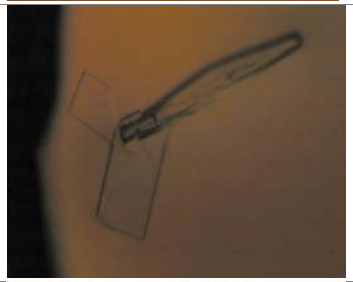
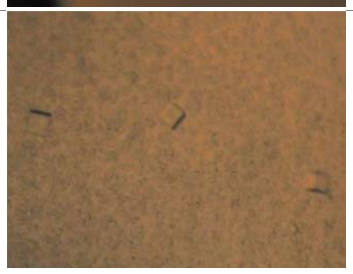
Although $Rbg1_{fl}$-$Tma46_{137-345}$, $Rbg1_{fl}$-$Tma46_{154-345}$ and $Rbg1_{fl}$-$Tma46_{205-345}$ were all used for the crystallization trials, only the complex $Rbg1_{fl}$-$Tma46_{205-345}$ could be successfully crystallized. The representative purification result for the $Rbg1_{fl}$-$Tma46_{205-345}$ construct is given in figure 4.1 with the chromatograms from the nickel chelating affinity (4.1.A) and size exclusion (4.1.B) chromatographic methods as all the other proteins were purified following the same protocol. Representative samples of the final purified proteins run on an SDS-PAGE gel is as shown in figure 4.1.C. Human Lerepo4 full length could neither be expressed in bacteria, but the C-terminal (220-396) could be obtained as shown in lane 8. The N-terminal $Lerepo4_{1-220}$ protein was obtained from Isabel Perez in the laboratory. Most of the expressed proteins, especially complexes, were highly soluble and could be concentrated to high concentrations even as much as 180 mg/ml. Exceptions were Rbg1 and Drg1 which had to be concentrated to much lower levels in order to avoid precipitation.

As previously reported, $Gir2_{fl}$ protein migrated at a size of around 70 kDa in an SDS-PAGE gel although the molecular weight of the protein is 31 kDa (lane 6). This was due to the anomalous electrophoretic behavior contributed mainly by its highly acidic N-terminal RWD domain (Alves and Castilho 2005, Alves *et al.,* 2004). The Rbg2 full length protein in complex with C-terminal dfrp fragment of Gir2 ($Rbg2_{fl}$-$Gir2_{174-238}$) was also obtained, the complex represented in lane 5, figure 4.1.C.

4.2 Crystallization trials of the $Rbg1_{fl}$-$Tma46_{205-345}$ (DRG-DFRP) complex

Of the various crystallization trials setup for the DRG-DFRP proteins, initial crystals were obtained for $Rbg1_{fl}$-$Tma46_{205-345}$ utilising the vapour diffusion sitting drop method. The crystals were small and grew at 21ºC using 40 mg/ml of more than 95% pure protein complex (as visualized in an SDS-PAGE gel). Some of the different conditions in which crystals were obtained are as given in table 4.1.

Table 4.1: Reservoir conditions in which initial $Rbg1_{fl}$-$Tma46_{205-345}$ crystals appeared. The crystals appeared in some of the commercial crystallization screens in 96-well plates.

S. No.	Image	Condition	Screen	Temp. (ºC)	No of Days
1.		4M Sodium formate	Hampton CS I & II	21	3-5
2.		0.1 M Tris Hydrochloride pH 8.5, 2.0 M Ammonium sulfate	Hampton CS I & II	21	5-7
3.		0.2 M Magnesium chloride, 0.1 M Bis-Tris pH 5.5, 25 % w/v PEG 3350	JCSG	21	5-7
4.		2 M Ammonium sulfate	Nextal Classics	21	5-7

In order to obtain larger crystals, these conditions were further optimized in 24-well plates. Particularly, large clustered rod-like crystals were obtained in 2.8 M Sodium formate at 21ºC in the 24 well plates which took around 3-5 days to grow. The crystals obtained in this condition were run in a gel after washing five times in reservoir solution. This confirmed the crystals to belong to the $Rbg1_{fl}$-$Tma46_{205-345}$ complex (Figure 4.2). Among the several cryoprotectant solutions tried (Glycerol, Ethylene glycol, Glucose, Sucrose, PEGs, PVPK15, Lithium acetate, Lithium fluoride, 2,3-Butanediol, Sodium formate in increasing concentrations etc.) the crystals were found to be stable only in around 50% Sodium malonate pH 7.0 or Paratone-N (Hampton Research). They diffracted poorly to around 3.1 Å and belonged to the primitive monoclinic lattice and were suggested to contain 4 molecules in the asymmetric unit (Mathews coefficient of 3.54, cell volume of 1615558, calculated solvent content of 65.3 %). The optimized crystals from the rest of the conditions as given in table 4.1 diffracted to a much lower resolution and were not followed further.

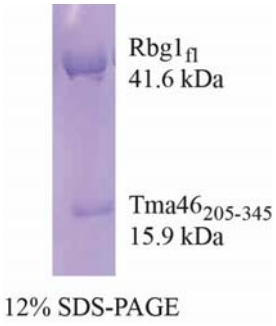


Figure 4.2: SDS-PAGE gel of $Rbg1_{fl}$-$Tma46_{205-345}$ protein from the crystals. The crystals obtained were washed and run on a polyacrylamide gel showing the presence of the two individual proteins $Rbg1_{fl}$ and $Tma46_{205-345}$.

Interestingly, the primitive monoclinic crystals (a=84.3 b=86.1 c=224.4 α= 90º β=97.2º γ=90º) obtained when dehydrated stepwise (in 1% increments; at the BM14 beamline, ESRF) from its initial relative humidity of 91% to a humidity of 78%, could be indexed as either primitive tetragonal (a=b=~86, c=~222, α=β=γ=90º) or orthorhombic (a=~85, b=~86, c=~222, α=β=γ=90º). This change in lattice was seen to be reproducible and reversible, i.e. changing from tetragonal/orthorhombic back to monoclinic when the relative humidity increased back towards 91%. Unfortunately, no improvement in the diffraction quality was observed for the dehydrated crystals, the resolution being lower than that of the starting monoclinic crystals. Also, integration and scaling of the images from the dehydrated crystal was not always successful due to the low quality of diffraction.

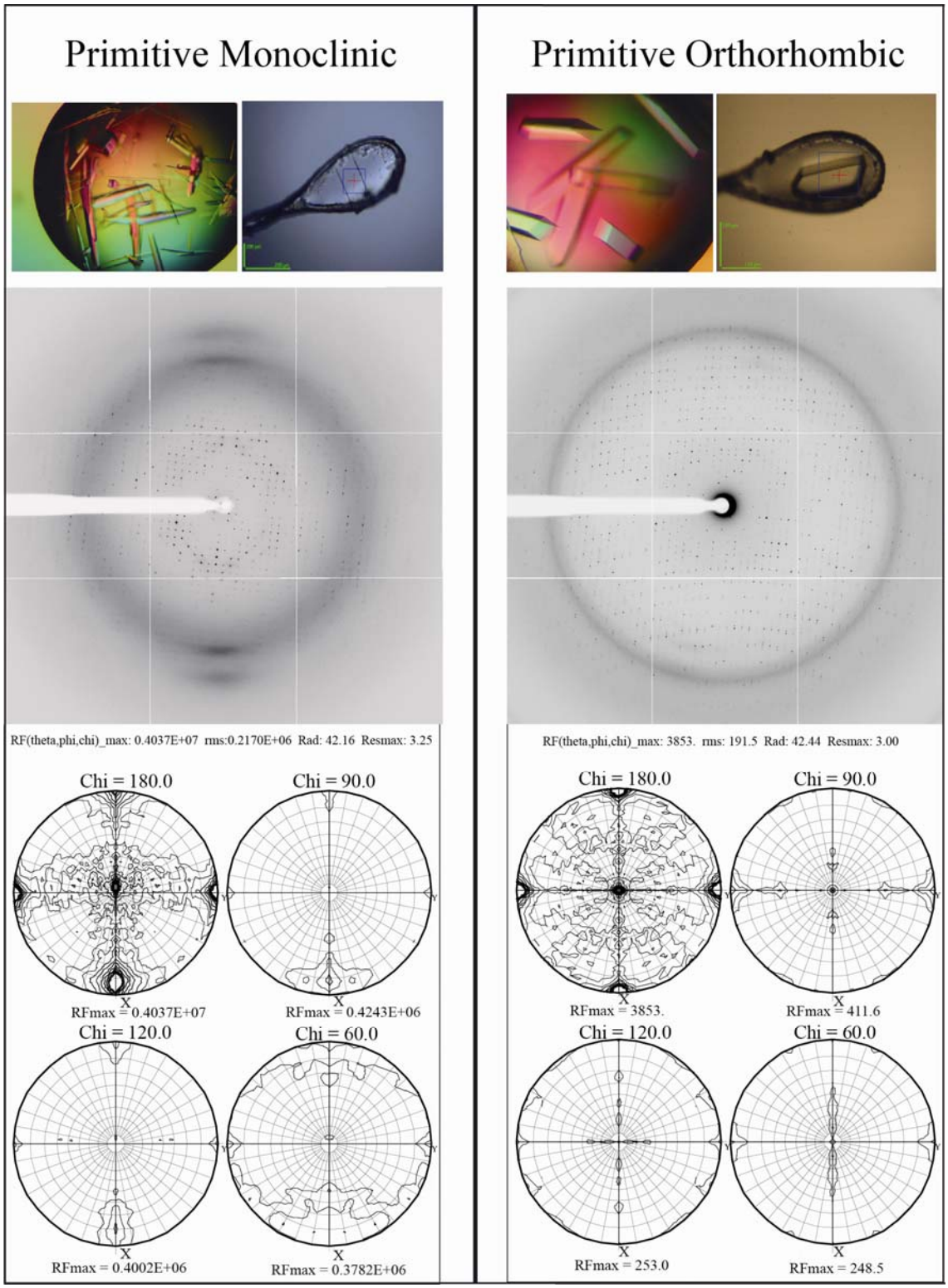


Figure 4.3: Comparison of the primitive monoclinic and orthorhombic crystals. The crystals belonging to the two lattices and their diffraction pattern is given. The self-rotation functions are also shown.

Later on, heterogeneous primitive monoclinic and orthorhombic crystals were also obtained by optimizing the crystallization conditions i.e. by performing a 96-well screen containing 85% of a solution of home-made 2.8 M Sodium formate with 15% of the conditions of Crystal screen I & II from Hampton which came up with Tri-sodium citrate dihydrate pH 6.5 as an additive. Large, diffraction quality crystals of typical dimensions 0.3 mm x 0.1 mm x 0.05 mm were further obtained by setting up 24-well plates at 4 ºC in the condition 2.38 M Sodium formate, 0.2-0.25 M Tri-sodium citrate dihydrate pH 6.5 which took about 2 weeks to grow. These crystals were found to be belonging to the primitive orthorhombic lattice (a=84.89 b=86.2 c=224.89 α= 90º β=90º γ=90º). A comparison of the crystals belonging to the two lattices (monoclinic and orthorhombic), along with the diffraction pattern and self-rotation functions are given in figure 4.3.

As can be seen in the self-rotation function graph for the primitive monoclinic kind of crystals, apart from the crystallographic two-fold axis along the y direction, a peak is seen in the xz plane which is shifted around 7º towards the z axis. This shows that the crystal has a near (pseudo) orthorhombic symmetry. The self-rotation function for the primitive orthorhombic crystal on the other hand showed peaks for the three binary axes indicating a clear orthorhombic symmetry. The orthorhombic crystals seemed to contain two molecules in the asymmetric unit (Mathews coefficient of 3.55, cell volume of 1634599.625 $Å^3$, calculated solvent content of 65.41%). The data for the orthorhombic crystals showed systematic absences in k and l reflections suggesting $P22_12_1$ space group. Since the shorter 2_1 screw axis comes first according to the International Union of Crystallography (IUCr) convention, the h,k,l indices were permuted to k,l,h using pointless (ccp4i) giving space group $P2_12_12$.

4.3 Structure determination of the yeast $Rbg1_{fl}$-$Tma46_{205-345}$ (DRG-DFRP) complex

Molecular replacement trials using homologous structures to the Rbg1 G-domain, like BsObg (PDB id: 1LNZ), TtObg (PDB id: 1UDX), EngA (PDB id: 2DYK), FeOB (PDB id: 3A1W) and HsYchF (PDB id: 2OHF) with 39, 34, 30, 29, 28 % identity respectively in the G-domain, and human Drg1 TGS NMR solution structure (PDB id: 2EKI) with 65% identity to Rbg1 TGS domain came out to be unsuccessful. The reason could be that Rbg1 G-domain makes up to only 30% of the complex structure and the TGS domain to around 15%. Homologous structures to Tma46 were also not available. Hence only 45% of the asymmetric unit (containing two heterodimers) is covered by the available homology models.

Hence, in order to obtain the phases for the structure solution, several soaking and co-crystallization trials with heavy metals (Hampton Heavy metal screen (which include Hg, Au, Pt, W, Sm, La, Eu, Gd, Lu, Yb, Dy, Pr, Nd, Ho and Re salts), $Ta_6Br_{12}^{2+}$, Rubidium, Barium, Caesium, halides (Br, I) etc.) were tried, some selected based on prediction from

HATODAS (Heavy-atom Database System, (Sugahara *et al.,* 2005) and protein band shift in native gel electrophoresis (Figure 4.4). Many of the heavy metals seemed to ligate the protein as evidenced in the figure. Data sets of around 3.0 Å resolution were collected from some of the derivatized crystals like Hg, Pt, W, Re, Yb and Ta which showed clear peaks at their L-absorption edge in X-ray fluorescence spectra. However, none of the derivatives gave sufficient isomorphous/anomalous differences and the limited and partial occupancy of the heavy atom sites located were not enough for experimental phasing.

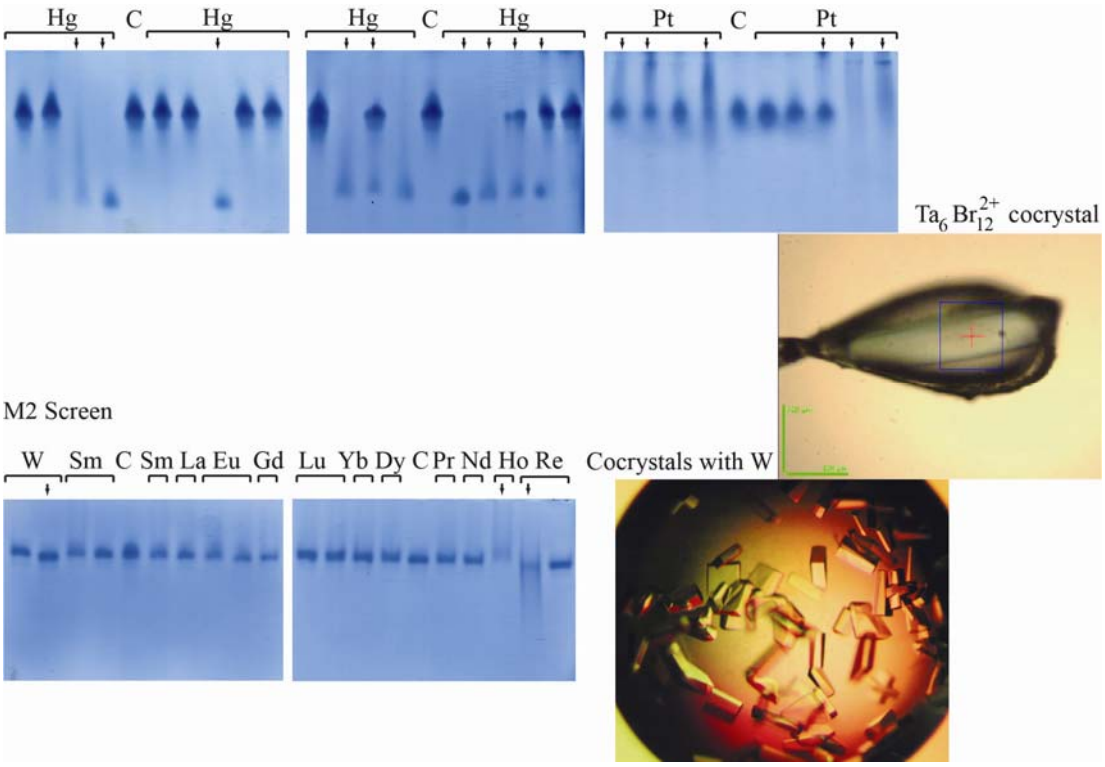


Figure 4.4: Heavy-atom derivatization. Native gel electrophoresis in a 7% polyacrylamide gel for detection of protein-heavy metal derivatization showed binding to some of the mercury, platinum, tungsten, ytterbium and rhenium salts as indicated by the arrows. The co-crystallized crystals obtained with Ammonium tetrathiotungstate and $Ta_6Br_{12}^{2+}$ are also given.

In parallel, selenomethionine incorporation into the $Rbg1_{fl}$-$Tma46_{205-345}$ protein was achieved by the autoinduction method and $_{SelMet}Rbg1_{fl}$-$Tma46_{205-345}$ was found to crystallize in the same reservoir conditions as the native protein. $Rbg1_{fl}$-$Tma46_{205-345}$ complex contains around 10 potential selenomethionine sites for which phase solution was difficult to obtain from the initial P2 kind of crystals (3.1 Å, 4 heterodimers per asymmetric unit). For the $P2_12_12$ crystals (2 heterodimers per asymmetric unit), although various phasing techniques were used (SAD, MAD, SIR, SIRAS) using the peak, remote and inflection data collected from a single $_{SelMet}$Rbg1-$Tma46_{205-345}$ crystal, sufficient phasing signal could be obtained ultimately, only by the SIRAS method (Single wavelength Isomorphous Replacement and Anomalous Scattering).

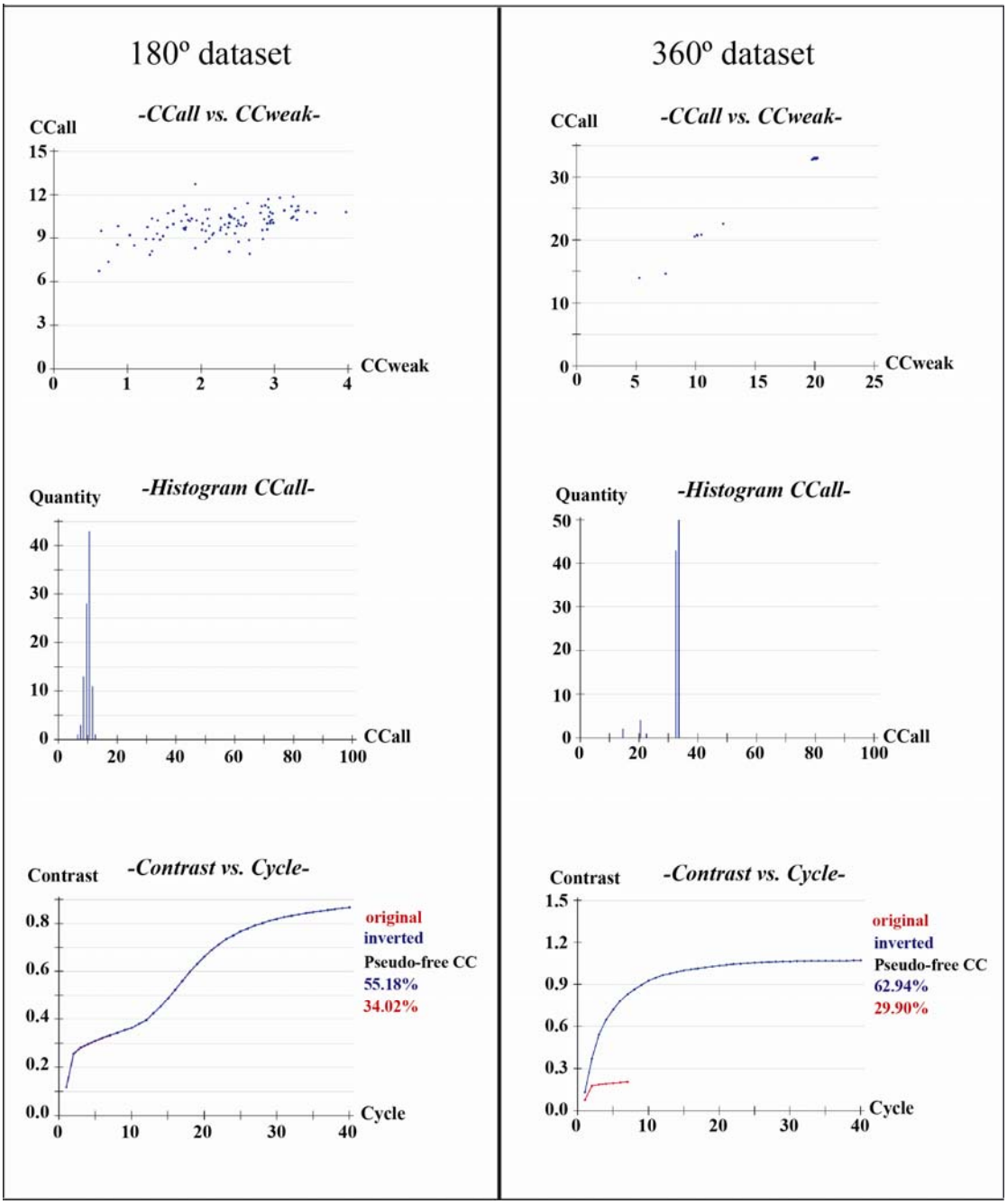


Figure 4.5: Comparison of the phasing signals from 180º and 360º dataset. The two datasets were collected from a single $_{SelMet}$Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ crystal at the peak wavelength.

The native dataset used for the SIRAS method was collected from a single crystal to a resolution of 2.82 Å. The dataset consisted of 360 images at a rotation angle (Δφ) of 0.5º, exposure of 1s at ID14-4 (ESRF) with a crystal to detector distance of 385.94 mm. Two datasets, one rotating 180º and the other 360º (Δφ = 0.5º, exposure time = 1s, resolution = 2.88 Å, crystal to detector distance of 444.48 mm), at the peak wavelength (λ = 0.9794 Å)

for selenium were also collected from a selenomethionine-derived crystal. Analysis of both the datasets showed that the localization of the heavy atom sites was only possible from the 360º complete data set as shown in figure 4.5. The CC_{all} graph from hkl2map shows that compared to the 180º dataset, the 360º dataset showed more than three heavy atom sites with good CC_{all}/CC_{weak}. This shows the importance of having a dataset with a higher redundancy which improved the accuracy of the anomalous signal sufficiently enough to obtain better heavy atom locations, a consequence of which resulted in good phases.

The data collection statistics for the native and SelMet peak 360º dataset are as given in table 4.2.

Table 4.2 Data collection statistics of $Rbg1_{fl}$-$Tma46_{205-345}$.

Data Collection			
Space group	**P2**	**$P2_12_12$**	
Unit cell parameters			
a, b, c (Å)	84.32, 86.06, 224.39	86.2, 224.89, 84.89	
α, β, λ (º)	90.0, 97.2, 90.0	90.0, 90.0, 90.0	
	Native	**Native**	**SeMet (peak)**
No of crystals used	1	1	1
Wavelength (Å)	0.9795	1.0332	0.9795
Resolution (Å)	111.32 – 3.00	79.42 - 2.67	80.38 - 2.88
No. of observations	208540	340325	527651
No. of unique Reflections	63656	47671	37866
Completeness (%)	99.2 (99.2)	99.9 (99.9)	99.6 (99.6)
Multiplicity	3.28	7.14	13.93
Mean $I/\sigma(I)$ [a]	8.1 (1.8)	15.6 (4.4)	21.6 (5.9)
R_{meas} (%) [b]	9.8 (72.6)	8.1 (43.4)	8.3 (42.3)
Average mosaicity	0.61	0.38	0.48

Numbers in parenthesis indicate the highest resolution shell statistics.

[a] Mean $[I/\sigma(I)]$ is the average of the relation between the intensity of the diffraction and the background.

[b] $R_{meas} = \{\Sigma_{hkl}\,[N/(N-1)]^{1/2}\,\Sigma_i\,|I_i(hkl) - \langle I(hkl)\rangle|\}\,/\,\Sigma_{hkl}\,\Sigma_i\,I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl)\rangle$ are the average intensities and N is the multiplicity of reflection *hkl*.

The structure solution was finally obtained using the automated program: autoSHARP. A search for 20 Selenium sites done in autoSHARP using SHELXC/SHELXD resulted in an initial set of 9 sites. This increased to 16 selenium sites after the cycles of heavy atom refinement and phasing, and analysis of the residual maps. The phasing statistics from the final SHARP run is as given in table 4.3. The overall phasing power of the anomalous differences dropped below 1 at 3.45 Å. For the isomorphous differences, the phasing power dropped below 1 at 6.88 Å for acentric and 79.42 Å for centric reflections. Initial phases obtained were then improved by density modification selecting the inverted hand (Score of 0.560 (2 molecules/AU) vs. 0.196 for the original hand) and a final solvent flattening using 57.8% solvent content giving phases with an overall correlation on E^2/contrast of 3.9208. 32 chains with 599 residues with a connectivity index of 0.87 were built by the final automated model building step. 13 of these chains comprising 406 residues were docked into the sequence (R_{factor}/R_{free} - 0.262/0.376) in 140 cycles. The rest of the model building was done manually in coot taking into account the remaining selenium sites, followed by the refinement cycles.

Table 4.3 Phasing statistics of Rbg1$_{fl}$-Tma46$_{205-345}$.

Phasing	
Overall Phasing power	
Anomalous differences	1.644
Centric/acentric (Isomorphous differences)	0.609/0.705
Overall figure-of-merit	
Centric/acentric	0.146/0.305

During the initial refinement cycles, non-crystallographic symmetry (NCS) restraints were maintained between molecules of Rbg1 and molecules of Tma46 in the asymmetric unit, using tight or medium main chain with medium or loose side chain. The NCS restraints were released during the final stages. TLS groups were defined for anisotropic refinement. The fine-tuned final 17 groups selected for the TLS refinement comprised of Rbg1 (chain A 2-45, 53-125/131-174/233-299, 175-232, 300-369; chain B 2-53, 54-91/98-125/133-174/233-299, 175-232, 300-369) and Tma46 (chain C 214-240, 241-267, 268-282, 302-313, 314-338; chain D 214-240, 241-267, 268-282, 320-336). The weighting of the X-ray and geometry terms was fixed to 0.05 as this seemed to give better geometry parameters and lower R factor values. The refinement statistics are as given in table 4.4.

Table 4.4 Data refinement statistics of $Rbg1_{fl}$-$Tma46_{205-345}$.

Refinement	
Resolution (Å)	56.22-2.67 (2.74-2.67)
No. of reflections R_{work}	46457
Reflections used in R_{free}	1200
R_{work} [a]	19.7%
R_{free} [b]	22.2%
Mean B-factor-Overall	76.935
Stereochemistry	
Res. in favoured regions (%)	89.7
Res. in allowed regions (%)	10.2
Number of atoms	
Protein	7030
Water	77
Rmsd [c]	
Bond lengths (Å)	0.007
Bond angles (º)	1.233
Residues modeled	
Rbg1 A	2-45, 53-125, 131-369
Rbg1 B	2-91, 98-125, 133-369
Tma46 C	214-282, 302-338
Tma46 D	214-282, 320-336
Residues with missing side chain	
Rbg1 A	Lys329
Rbg1 B	Ala46, Ser47, Ser48, Ser50, Lys369
Tma46 C	Glu307
Tma46 D	Leu214, Glu215, Asp320
PDB id	**4a9a**

[a] $R_{work} = \Sigma_{hkl}$ {[$F_{obs}(hkl)$] - [$F_{calc}(hkl)$]} / Σ_{hkl} [$F_{obs}(hkl)$], where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.

[b] R_{free} corresponds to R_{work} calculated using 2.5 % of the total reflections selected randomly and excluded during refinement.

[c] *Rmsd* is the root mean square deviation.

The structure of $Rbg1_{fl}$-$Tma46_{205-345}$ complex was refined to 2.67 Å (R factor – 19.7%, R_{free} – 22.2%) and includes two molecules of the complex in the asymmetric unit although the protein elutes as a single heterodimer by size exclusion chromatography. The structure is given as molecule A and B (Rbg1) interacting with molecule C and D (Tma46)

respectively. In order to determine the orientation of the components of the asymmetric unit, an analysis of the interfaces was carried out using the PISA server which calculates interaction interfaces, surfaces and assemblies for macromolecular structures.

Table 4.5: Interface areas for $Rbg1_{fl}$-$Tma46_{205-345}$ in the crystal. The top interaction interfaces suggested by the Pisa server along with their area are given.

S. No.	Molecule I	Molecule II	Symmetry operation	Interface area Å^2	Δ^iG kcal/mol	N_{HB}	N_{SB}	CSS
1	C	A	x,y,z	2950.1	-42.5	29	16	1.000
2	D	B	x,y,z	2285.5	-34.1	21	7	1.000
3	B	A	x,y,z	824.7	-3.8	12	0	0.000
4	C	D	x-1,y,z	467.2	-8.2	0	0	0.000

As can be seen in table 4.5 and figure 4.6.A, molecules AC and BD (Rbg1-Tma46) gives a high value of interface area as obviously they form the complex. In the asymmetric unit however, the interface between the two heterodimers AC and BD could either be between molecules of Rbg1 A and B or between the Tma46 molecules C and D as can be seen in figure 4.6.B.

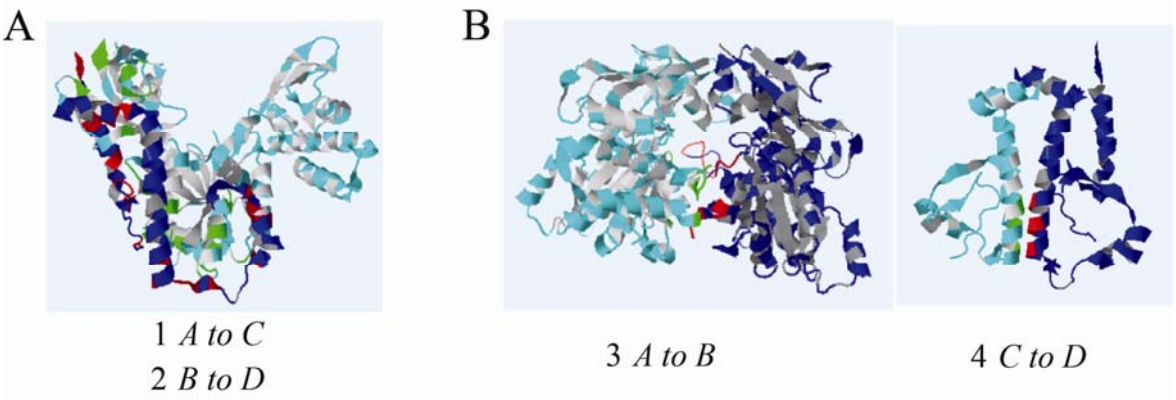


Figure 4.6: Interaction interfaces of the $Rbg1_{fl}$-$Tma46_{205-345}$ structure in the crystal.

A. $Rbg1_{fl}$-$Tma46_{205-345}$ (heterodimers AC and BD corresponding to 1 and 2 in table 4.5) complex has the highest interface area among those suggested by the Pisa webserver. The residues in contact are represented as red/green.

B. The interface area between $Rbg1_{fl}$ molecules A and B (3 in table 4.5) is almost double that of the interface between $Tma46_{205-345}$ molecules C and D (4 in table 4.5).

Pisa analysis suggests the next largest interface area, after that of the complexes, to be formed between molecule A and B with a value of 824.7 Å^2 followed by the interface between the α2 helix of molecules C and D (Tma46). However, the latter interface area was

only about half of that between molecule A and B (467.2 $Å^2$) and so this was taken as the interface between the asymmetric unit components.

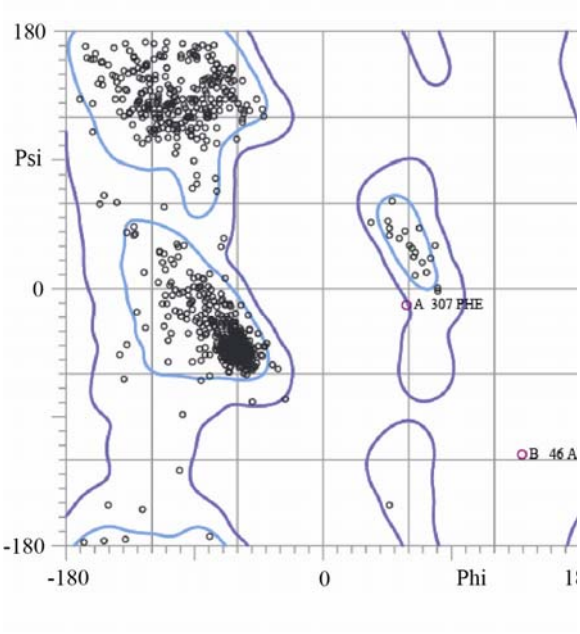


Figure 4.7: Ramachandran´s plot for the $Rbg1_{fl}$-$Tma46_{205-345}$ complex structure. The Ramachandran´s plot was obtained using the Molprobity webtool. The outliers A 307Phe ((phi: 58.7, psi: -11.8) and B 46Ala (Phi: 139.3, Psi: -116.0) seen in the plot were in regions of poor electron density.

The two heterodimers AC and BD superimpose with an RMSD of 1.8 Å over all Cα atoms. Residues that could be modelled for the protein complex are given in table 4.4. While 96.5% (molecule A) and 96.2% (molecule B) of the residues of Rbg1 could be modelled, that for Tma46 was far less, being 75.2% and 61% for molecule C and D respectively. Notably, some of the loops comprising the G-motifs in Rbg1 had poor electron density due to their flexibility. So was the case with the glycine-rich region connecting the N-terminal first 45 residues to the G-domain. Tma46 has a large segment (residues 283-301 in molecule C and 283-319 in molecule D) for which electron density was not seen. 77 water molecules in the first solvation shell were included. 95.7% of all residues were in the favoured regions of the Ramachandran plot and 99.6% were in the allowed regions according to Molprobity (Figure 4.7, (Davis *et al.,* 2007, Lovell *et al.,* 2003)). The coordinates and structure factors were deposited in the RCSB Protein Data Bank (PDB) with the id 4a9a.

4.4 Structural analysis of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex

The Rbg1 structure shows a multi-modular domain organization which included the well-conserved G-domain (G1+G2+G3 = 64-169; G4+G5 = 245-293) and the C-terminal TGS domain (294-369) which were reported previously. Additionally Rbg1 was found to contain two more domains: an N-terminal helix-turn-helix (HTH) subdomain (1-44) which lies adjacent to another subdomain formed by a 65 residue long insertion (170-244) between G3 and G4 of the G-domain, and the TGS domain at the C-terminus (Figure 4.8).

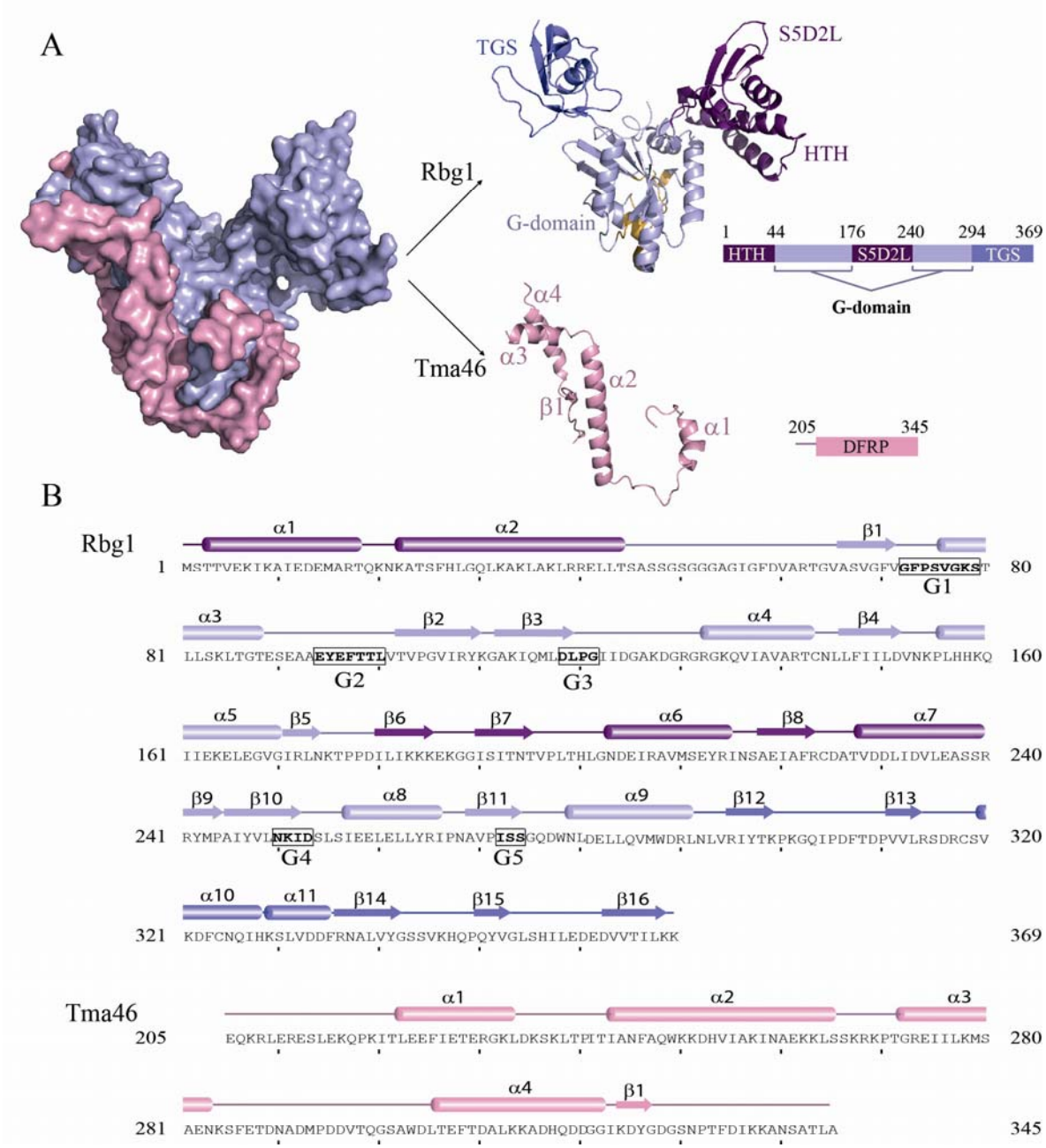


Figure 4.8: Structure of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex with sequence information.

A. Surface representation of the $Tma46_{205-345}$ C-terminal fragment (pink) enveloping $Rbg1_{fl}$ (pale blue) is shown on the left. The individual components are shown colour-coded on the right: Rbg1 with the different domains, G-domain (pale blue), HTH-S5D2L domain (purple), TGS domain (blue) and the Tma46 C-terminal fragment (pink). The GTP binding pocket is also represented with the five G-motifs coloured as orange.

B. The component sequences and secondary structure elements of the crystallized complex are represented with the G-motifs (G1-G5) given in boxes. Domain boundaries are indicated in the same colour scheme as in A.

Analysis of the $Rbg1_{fl}$-$Tma46_{205-345}$ structure shows that the C-terminal part of Tma46 is highly unstructured comprising only of helices which envelops over Rbg1 establishing contacts with its N-terminal G-domain and C-terminal TGS domain. The structural description which follows is based on $Rbg1_{fl}$-$Tma46_{205-345}$ molecule AC unless otherwise specified.

4.4.1 The Rbg1 G-domain

The G-domain of Rbg1 contains five α-helices (α3, α4, α5, α8 and α9) and six β-strands which forms a beta sheet structure (β1, β2, β3, β4, β10 and β11). The G1 motif (P-loop/Walker A) is located on the loop connecting β1 to α3, G2 (Switch I) in between α3 and β2, G3 (Switch II/Walker B) at the end of β3 strand, G4 in β10 and G5 motif in β11. The G-domain of the two molecules A and B of Rbg1 are very similar to each other except in the loops containing the five G-motifs. RMSD between the two Rbg1 molecules (chain A & B) in the asymmetric unit considering only the main chain atoms of the modelled residues was 1.38 Å. By superimposing human OLA1 structure (Obg-YchF family member) in complex with AMPPCP (PDB id: 2OHF) onto the G-domain of the two chains, it was seen in comparison that the conformation of the P-loop in chain A and chain B of Rbg1 were not exactly the same. The loops containing the G2 and G3 motifs were also shifted although the electron density in this region was poor.

4.4.2 The Rbg1 TGS domain

The C-terminal TGS domain has predominantly beta sheet structure with five beta strands (β12-β16) and two helices (α10 and α11), and is made up of about 65 amino acids from residues 294 to 369 (figure 4.8). It was found to be quite similar to the previously reported NMR solution structure of human Drg1 TGS domain (PDB id: 2EKI) with an RMSD of 1.31 Å superimposing the main chain atoms. The structures which come up in a DALI search with the TGS domain include alignments with the Threonyl tRNA synthetase (Z-score – 7.4, RMSD of 1.6 Å), YchF hOLA1 TGS domain (Z-score – 7.1, RMSD of 2.4 Å) and ubiquitin (Z-score – 6.1, RMSD of 2.5 Å).

4.4.3 The Rbg1 S5D2L domain

Previous reports have shown that DRG factors have an insert sequence between the G3 and G4 motifs of the G-domain which was not found in any of the other Obg family members and had no sequence homology to known domains (Wout *et al.,* 2009). From the structure it was seen that this region (residues 176-240) folds as an independent domain emerging from the G-domain and a short parallel beta sheet formed by β5 and β9 connects this domain between the segments containing the G3 and G4 motifs.

A sequence search of the domain found no results in PDB databank. Structural similarity search with DALI however, came up with the 30S ribosomal S5 subunit protein C-terminal domain which was the most similar, aligning with an RMSD of 2.2 Å over 51 residues even though the sequence identity was very low (12%). DALI search suggested that the topology of this domain was related to that of the "Ribosomal protein S5 domain 2-like" superfamily. Hence we named this novel domain in Rbg1 as the S5 Domain 2-Like (S5D2L) domain. The aforementioned superfamily has around 14 members as according to Pfam (Clan - CL0329). The superfamily is also annotated in InterPro database (IPR020568) as well as CATH (3.30.230) and SCOP (54211) classifications. The members belonging to the "Ribosomal protein S5 domain 2-like" superfamily include the 30S/40S ribosomal S5 subunit C-terminal domain, Elongation Factor-G domain IV, DNA mismatch repair protein C-terminal domain, Topoisomerase VI B subunit, Ribonuclease P and PH, Lon protease C-terminal domain, UDP-3-O-acyl N-acetylglycosamine deacetylase protein (LpxC), Imidazole glycerol-phosphate dehydratase (IGPD), GHMP kinases, Formaldehyde-activating enzyme (FAE) among others.

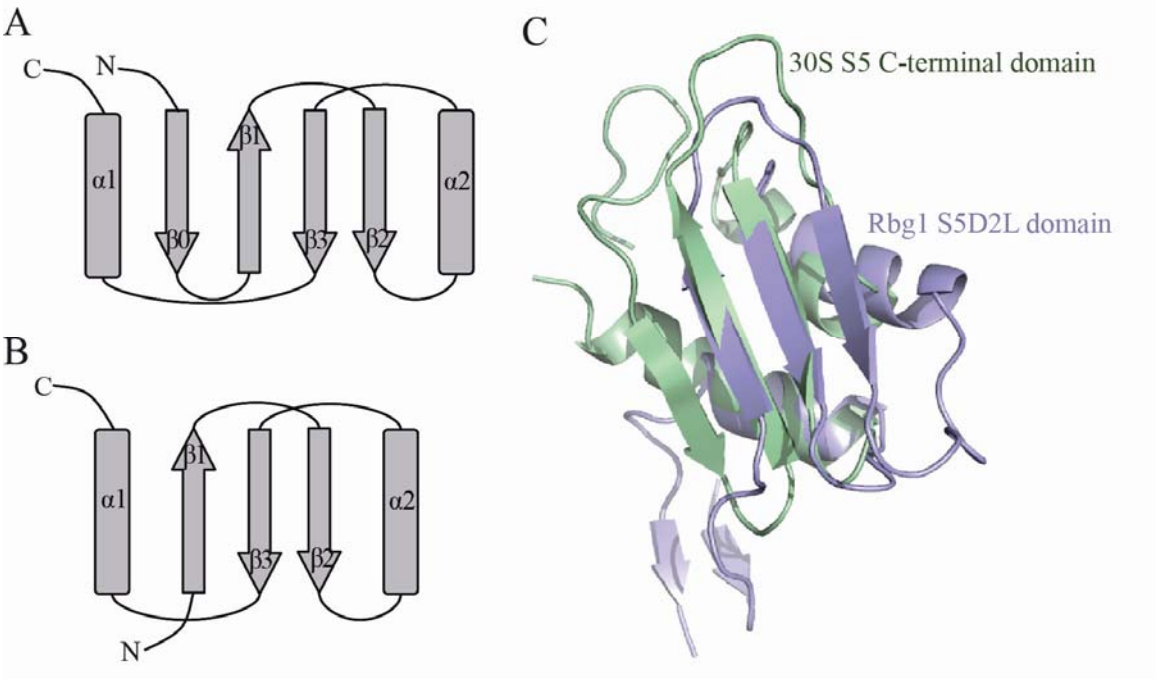


Figure 4.9: Architecture of the S5D2L domain

A. The unusual βββαβα fold found in many of the 14 member Ribosomal protein S5 domain 2-like superfamily which include the 30S and 40S S5 subunit protein C-terminal domain, Elongation factor domain IV, Mevalonate pyrophosphate decarboxylase (GHMP kinase family member) etc.

B. In comparison, the S5D2L domain of Rbg1 has a ββαβα fold lacking the N-terminal beta strand (β0).

C. Rbg1 S5D2L domain (blue) superposed over bacterial ribosomal S5 C-terminal domain (green) sharing the same topology.

Structural alignment shows that whereas the members of this superfamily have a βββαβα fold, S5D2L domain has a ββαβα fold lacking the first β strand (Figure 4.9). The residues Gly and Arg of the 30S ribosomal S5 subunit protein known to cause ribosomal ambiguity when mutated, were seen to be fully conserved in Rbg1 (Gly189 and Arg207). Equivalent residues are also present in Elongation factor G (EF-G) domain IV. The nature of these residues is however not universally conserved, as they are not found in the GHMP kinase family members.

4.4.4 The Rbg1 HTH domain

The N-terminal of Rbg1 comprises of two amphipathic helices, α1 and α2 (2-44) which form a previously unnoticed Helix-Turn-Helix (HTH) subdomain which lies adjacent to the S5D2L domain. The HTH is stabilized mainly by a hydrophobic zipper between the two helices (Figure 4.10), five leucines positioned 3-4 residues apart in the longer helix α2 contributing to the zipper while the other side of the helix α2 makes both hydrophilic and hydrophobic interactions with the S5D2L helices α6 and α7. Additionally, the interface between α1 and α2 is also stabilized by hydrogen bonds between Glu12 (α1) - Arg39 (α2) and Glu14 (α1) and His27 (α2).

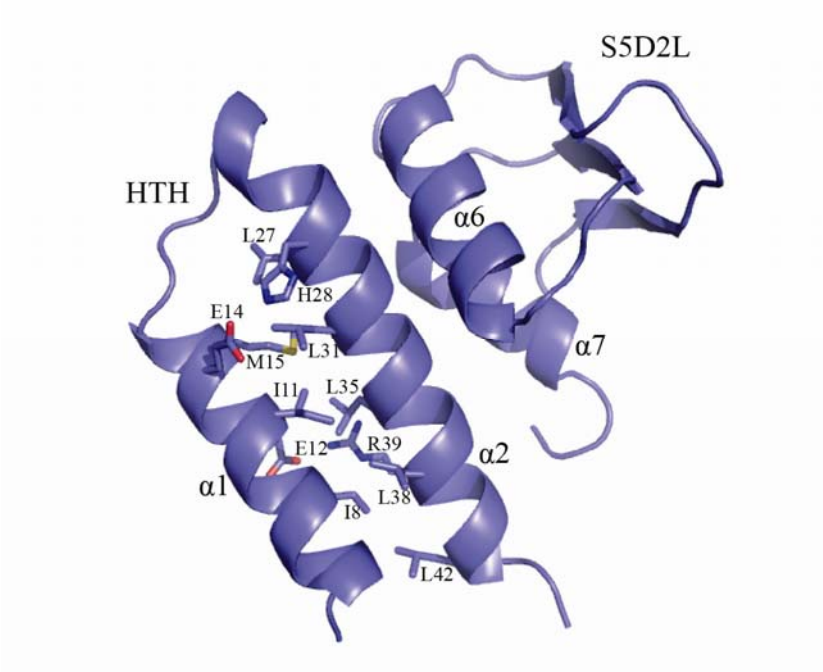


Figure 4.10: Structure of the HTH and S5D2L domains. The interacting residues between α1 and α2 helices forming the HTH domain are shown along with the adjacent S5D2L domain.

The TGS, HTH and S5D2L domains lie on the distal part of the GTP binding pocket (Figure 4.8). Electrostatic surface potential analysis shows an extensive positively charged

surface formed partly by the TGS, HTH, S5D2L and part of the G-domain opposite to the GTP binding site (Figure 4.11)

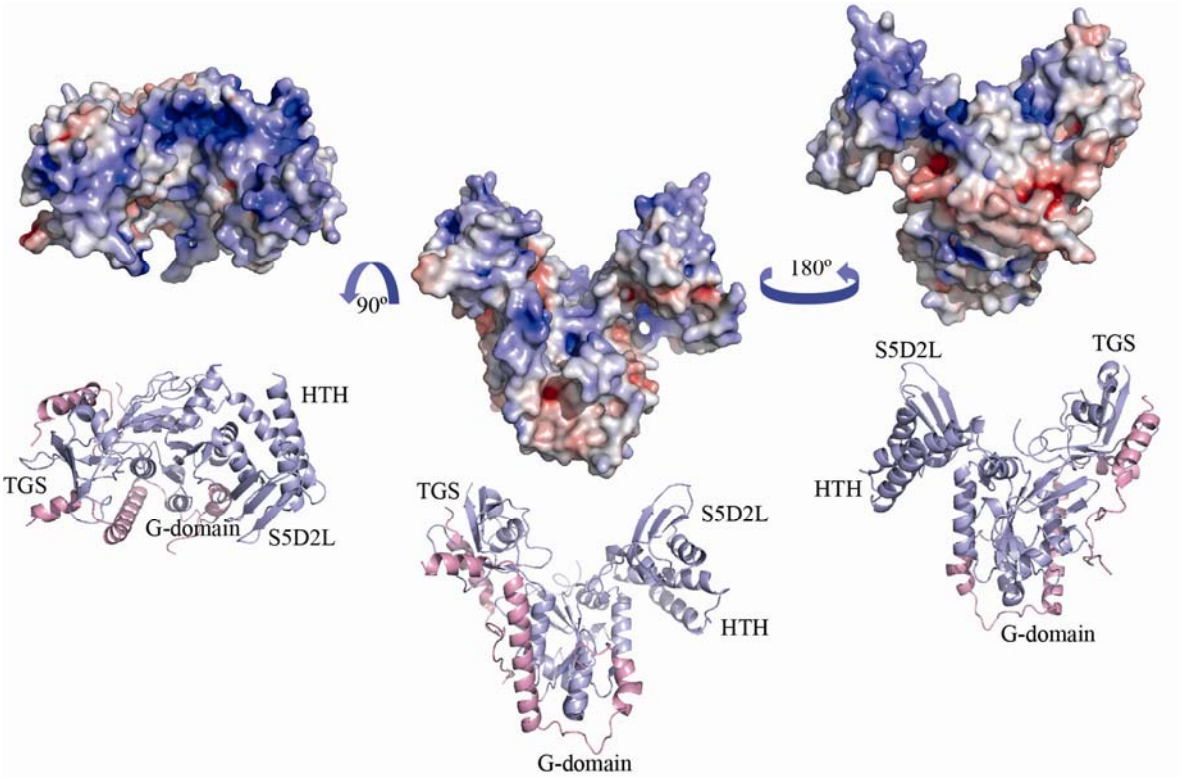


Figure 4.11: Electrostatic Potential Surface.

The solvent-accessible surface electrostatic potential of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex as calculated by APBS (Pymol) is shown as a surface alongside the cartoon representation. The potential is given with the negative (red) and positive (blue) contour levels in the range from -8.0- +8.0 kBT respectively. The left figure shows the positively charged surface formed partly by the G-domain, TGS, HTH and S5D2L domains.

4.4.5 Tma46 structure and interaction with Rbg1

Tma46 fragment present in the structure shows a non-globular type of fold predominantly formed by alpha helices interconnected by coils. The residues of $Tma46_{205-345}$ that we could model into the electron density is comprised of mainly four helices (numbered in this study as α1- α4; note that around 19 residues were not modelled in between α3 and α4) and a short beta strand which forms beta sheet structure with adjacent beta strands from the G-domain of Rbg1 (β2, β3, β1, β4, β10 and β11). Tma46 helices envelops Rbg1 forming an extended and extensive interface over its G-domain and TGS domain with no contacts with the HTH, S5D2L domains (Figure 4.8).

The total buried surface area upon binding of $Tma46_{205-345}$ to Rbg1 (molecule AC) is 2950 $Å^2$ with about 86 residues from Rbg1 (molecule A) interacting with 67 residues of $Tma46_{205-345}$ (molecule C) (Table 4.5). Based on weak sequence similarities between Dfrp1 and Dfrp2, it was earlier proposed that the region corresponding to residues 280 to 332 of Tma46 constituted a DFRP domain responsible for interaction with Rbg1 (Ishikawa *et al.,* 2005). Based on two-hybrid screen results, it was also previously suggested that Tma46 residues 254-296 could be sufficient for binding to Rbg1 (Wout *et al.,* 2009). The complex structure demonstrates that the region of Tma46 contacting Rbg1 is larger than these earlier estimates and encompasses residues 216-279 and 302-338. This observation suggests further that the DFRP region defined earlier is only a fraction of the biologically relevant unit involved in DFRP-DRG protein interaction and that the region of Gir2/Dfrp2 involved in contacting Rbg2 could also be larger.

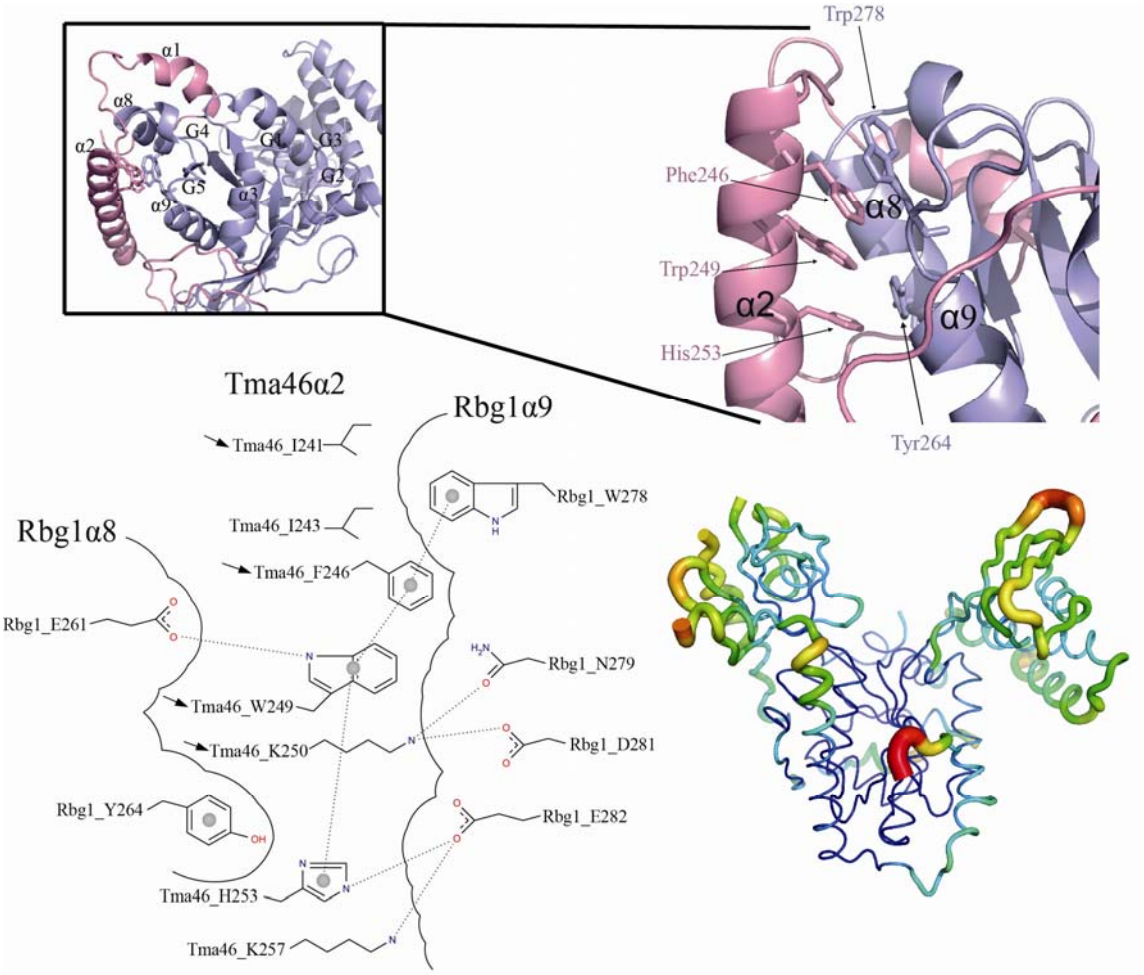


Figure 4.12: Interface between Rbg1 and Tma46 in the region of pi stacking adjacent to the G-domain. Inset showing the GTP binding pocket (G1-G5 motifs) of Rbg1 with its α8 helix inserted between Tma46 helices α1 and α2. Zoom in on the residues involved in pi stacking interactions in this area is shown. Also given is a 2D representation of the interface. The B factor putty representation (Pymol) is given for the complex structure.

One of the most prominent interacting surfaces between Rbg1 and Tma46 comprises α8 helix of Rbg1 (between G4 and G5 motifs), which inserts between helices α1 and α2 of Tma46 themselves interacting in turn also with Rbg1 α5 and α9 respectively. Helix α2 is longest in the fragment of Tma46 solved, and extends from the Rbg1 G-domain to the TGS domain that it contacts with the loop just after its C-terminal end. Interestingly, the interface between Tma46 α2 and Rbg1 α8/α9 consists mainly of aromatic ring containing residues, which form a pi stacking interaction. Specifically, Tyr264 (α8) and Trp278 (α9) of Rbg1 forms a stacking interaction with Phe246, Trp249 and His253 of α2 of Tma46 as given in figure 4.12.

4.5 Design, expression and purification of DRG and DFRP mutant proteins

Some mutants of the DRG and DFRP proteins were designed for expression in *E.coli* for the purpose of performing further experimental studies. This included domain deletion mutants of the DRG proteins like the yeast Rbg1 and Rbg2 proteins lacking the TGS domain both in complex with DFRP C-terminal ($Rbg1_{1-294}$-$Tma46_{205-345}$, $Rbg2_{1-294}$-$Gir2_{174-238}$) and Rbg1 TGS domain in complex with two C-terminal truncated versions of Tma46 ($Rbg1_{272-369}$-$Tma46_{205-345}$ and $Rbg1_{272-369}$-$Tma46_{154-345}$). Drg1 TGS domain ($Drg1_{289-367}$) was obtained from Mercedes Spinola in the laboratory.

Expression vectors containing DRG S5D2L domain ($Rbg1_{176-240}$, $Rbg2_{174-240}$, $Drg1_{175-238}$) on its own was obtained by cloning from the corresponding full length protein constructs. A construct of Rbg1-Tma46 complex lacking the S5D2L ($Rbg1_{\Delta 175-243+G}$-$Tma46_{205-345}$) was also designed using the information obtained from the crystallographic structure. Since the S5D2L domain protrudes out from the Rbg1 G-domain, amino acid residues 175 to 243 which form part of the S5D2L were deleted adding a glycine residue in its place to make up for the gap in between the G-domain. This construct could be successfully expressed in the bacterial system and purified. Rbg1 GTPase inactive mutants targeting residues in the G1 motif like GFPSVGK**N** and GFPSV**AMN** were designed for use in GTPase activity assays.

Mutations in Tma46 targeting the residues involved in pi-stacking and also some conserved residues present in the dfrp domain in both Tma46 and Gir2 were also designed. Also includes the Rbg1-Tma46 complex which lacks the Tma46 α1 helix. The plasmids obtained of all of these mutant proteins from different sources (given in table 3.1) were successfully expressed and purified following the same protocol as used for the wildtype proteins (Refer section 3.1.2). The representative purified protein samples are shown in an SDS-PAGE gel (figure 4.13).

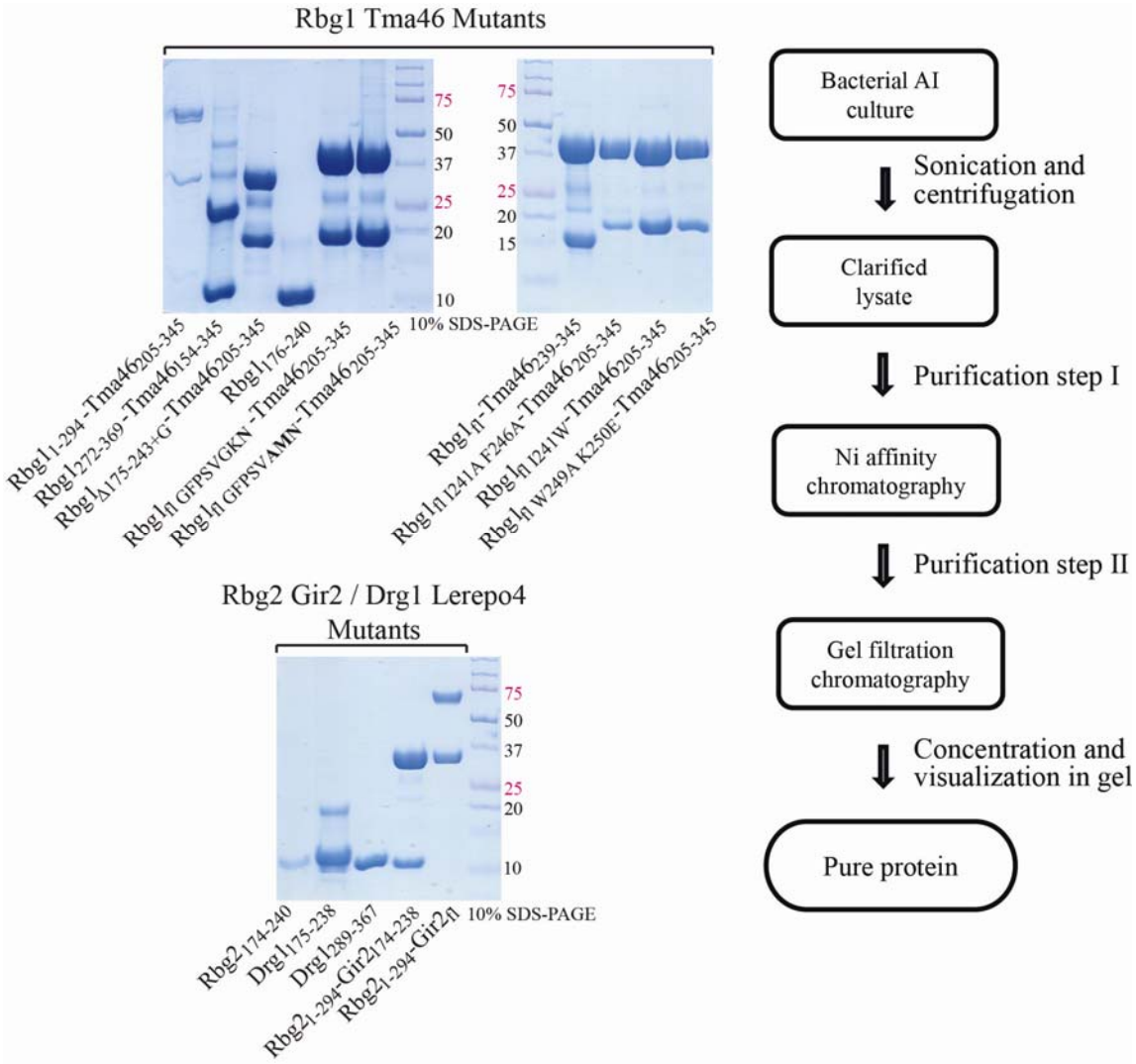


Figure 4.13: Purification results for the mutant proteins. SDS-PAGE gel showing the purified proteins is shown on the left while the general purification protocol followed is given on the right. The marker used for the protein gel was Precision Plus Protein Dual Colour Standards (BioRad).

A truncated form of Rbg1 containing only the TGS domain (272-369) still forms complex with Tma46 C-terminal fragment but shows weaker affinity in an imidazole gradient on a nickel chelating column (data not shown). Rbg1$_{1\text{-}294}$-Tma46$_{205\text{-}345}$ complex lacking the TGS domain also showed reduced Tma46 protein levels in purification steps suggesting that the overall interaction of Tma46$_{205\text{-}345}$ with both G-domain and TGS domain of Rbg1 is necessary for a complete stability of the binding. Additionally, the interaction between Rbg1 and Tma46 was not seen to be disrupted in the purification steps in the case of the Rbg1 protein lacking the S5D2L domain (Rbg1$_{\Delta 175\text{-}243+G}$-Tma46$_{205\text{-}345}$) expressing together with Tma46$_{205\text{-}345}$. This result further suggests that the Rbg1 S5D2L domain might not be contributing to any interaction with Tma46$_{205\text{-}345}$ as also seen from the crystallographic structure.

Mutants $Rbg1_{fl}$-$Tma46_{205\text{-}345\ G272A\ R273E}$ and, $Rbg1_{fl}$-$Tma46_{205\text{-}345\ W249A\ K250E}$, $Rbg1_{fl}$-$Tma46_{205\text{-}345\ I241A\ F246A}$ and $Rbg1_{fl}$ -$Tma46_{205\text{-}345\ I241W}$, the latter three targeting the pi stacking interaction region between Rbg1 α8 and Tma46 α2 did not disrupt the binding of Tma46 to Rbg1. Entire helix deletion of α1 in Tma46 also failed to disrupt the interaction between Rbg1 and Tma46 clearly due to the extensive surface of interaction between the two proteins.

Concomitantly, we designed a whole set of alpha helix deletions for Tma46 and domain deletions for Rbg1 for expression in yeast strains in the context of growth phenotype assays and immunoprecipitation studies performed in collaboration by the laboratory of Dr. Bertrand Séraphin. The outcome of these experiments is discussed in section 5.5. Tma46 helix deletions such as $Tma46_{1\text{-}320}$ (Δβ), $Tma46_{1\text{-}304}$ (Δα4-β), $Tma46_{1\text{-}268}$ (Δα3-β) and $Tma46_{1\text{-}238}$ (Δα2-β) from the C-terminal were designed as well as deletions from the N-terminal such as $Tma46_{\Delta 215\text{-}338}$ (Δα1), $Tma46_{\Delta 215\text{-}269}$ (Δα1-α2). An isolated alpha helix 2 mutant was also designed by converting residues 243 to 268 of Tma46 to alanines ($Tma46_{I243\text{-}R268\rightarrow 26A}$) in order to analyse the contribution of α2 specifically on the interaction to Rbg1 and its function. Residue mutants like $Tma46_{205\text{-}345\ G272A\ R273E}$, $Tma46_{205\text{-}345\ W249A\ K250E}$, $Tma46_{205\text{-}345\ I241A\ F246A}$ and $Tma46_{205\text{-}345\ I241W}$ were also designed for expression of the epitope tagged Tma46 mutant proteins.

HA tagged Rbg1 domain deletion mutants designed for expression in yeast included $Rbg1_{49\text{-}369}$ (ΔHTH), $Rbg1_{\Delta 175\text{-}243+G}$ (ΔS5D2L-1) and $Rbg1_{\Delta 171\text{-}239}$ (ΔS5D2L-2). The structural information provided a basis to study the effect of the individual components of the two proteins on the functionality of the Rbg1-Tma46 complex in yeast.

4.6 DRG proteins and GTPase activity

The presence of a core G-domain flanked by other domains and interacting with a structurally unstructured protein actuated our study of the GTP binding and hydrolytic activities of the DRG proteins in the context of its various structural elements.

4.6.1 GTP binding

The nucleotide binding of the DRG proteins was studied using an indirect method where the increase in protein unfolding temperature (Tm) upon binding to a ligand is measured. This thermal shift assay was performed here in the presence or absence of GTP, GDP or GTPγS. Nucleotide binding of the complexes $Rbg1_{fl}$-$Tma46_{205\text{-}345}$, $Rbg2_{fl}$-$Gir2_{174\text{-}238}$ and $Drg1_{fl}$-$Lerepo4_{220\text{-}426}$ was evidenced by an increased melting temperature in the presence of the GDP, GTP or non-hydrolysable GTPγS (Figure 4.14.A). Individual Drg1 also showed an increase in Tm suggesting GTP binding although we could not see the same for Rbg1 (Figure 4.14.B).

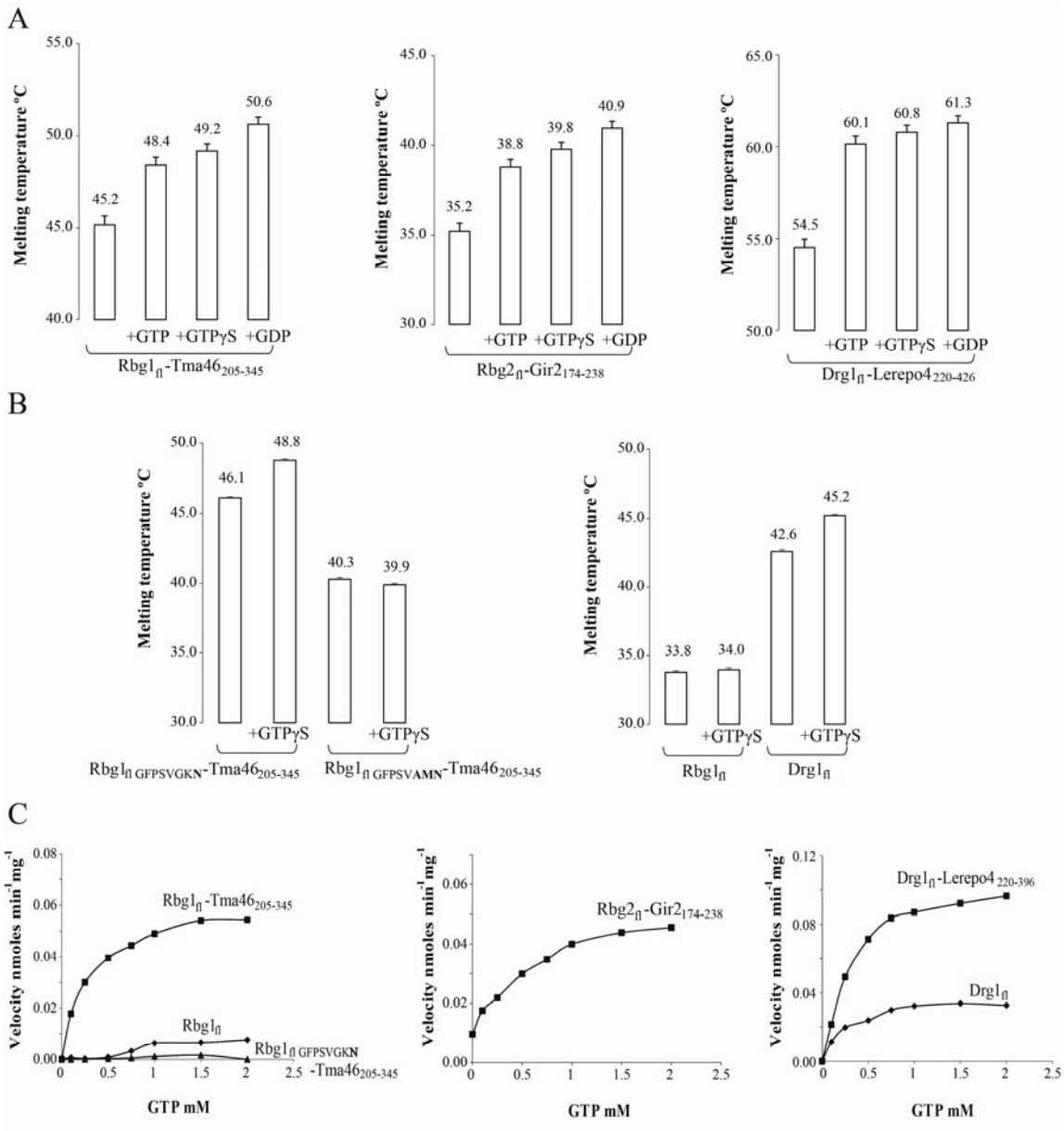


Figure 4.14: GTP binding and hydrolytic activity of the DRG-DFRP complex proteins. A. Presence of GTP, GTPγS or GDP (in 4 fold excess over the protein concentration) causes an increase in melting temperature of the $Rbg1_{fl}$-$Tma46_{205-345}$, $Rbg2_{fl}$-$Gir2_{174-238}$ and $Drg1_{fl}$-$Lerepo4_{220-426}$ complexes in the thermal shift assay indicative of nucleotide binding.

B. Also given are the thermal shift assay results for the catalytic site mutants and $Rbg1_{fl}$/$Drg1_{fl}$. Note also that Rbg1 and Drg1 melts at a much lower temperature than $Rbg1_{fl}$-$Tma46_{205-345}$ and $Drg1_{fl}$-$Lerepo4_{220-426}$ respectively.

C. GTP hydrolytic activity of $Rbg1_{fl}$-$Tma46_{205-345}$, $Rbg2_{fl}$-$Gir2_{174-238}$ and $Drg1_{fl}$-$Lerepo4_{220-426}$ are represented as a graph with increasing substrate concentration in the x axis.

Rbg1 protein containing mutations in the G1 motif (GFPSVGKS) were also analysed for GTP binding as negative control. Accordingly, two catalytic site mutants of $Rbg1_{fl}$, one involving the single mutation GFPSVGKS to GFPSVGK**N** ($Rbg_{fl\ GFPSVGKN}$) and the other a triple mutation, GFPSVGKS to GFPSV**AMN** ($Rbg_{fl\ GFPSVAMN}$) were tested both in complex with $Tma46_{205-345}$. The Rbg1 GFPSVGK**N** mutant was shown previously to be non-functional in yeast (Daugeron, 2011). This mutant however, showed around 2.5ºC increase in unfolding temperature in comparison to the wild type protein complex ($Rbg1_{fl}$-$Tma46_{205-345}$) as seen in figure 4.14.B suggesting some degree of GTP binding. For this reason, the Rbg1 GFPSV**AMN** mutant was designed as it was earlier reported for Mx proteins that mutating lysine in the G1 motif which binds to α- and β-phosphate of the guanine nucleotide to methionine abolished nucleotide binding (Pitossi *et al.,* 1993). As expected, the $Rbg1_{fl\ GFPSVAMN}$-$Tma46_{205-345}$ mutant showed no increase in melting temperature suggestive of lack of GTP binding as detected by the thermal shift assay. This experiment confirmed that the DRG proteins do bind GTP/GDP with the protein-ligand complex formation resulting in an increased stability of the protein as evidenced by the increase in unfolding temperature. We next analysed the ability of the DRG proteins to hydrolyse GTP *in vitro*.

4.6.2 GTP hydrolysis

GTP hydrolysis assays were performed with the overexpressed and purified DRG proteins to purity (homogeneity) of 95-99%. The assay was performed using a non-radioactive phosphate determination assay based on Malachite green reagent (Baykov *et al.,* 1988). In this assay, quantitation of the hydrolysis of GTP to GDP is assessed by the formation of a colorimetric complex between the inorganic phosphate generated and an ammonium molybdate-malachite green complex and the subsequent measurement of the absorbance at UV-visible range (Hoenig *et al.,* 1989, Lanzetta *et al.,* 1979).

Optimum conditions for the reaction were determined to be as follows: Buffer containing KCl (50 mM Tris pH 7.5, 100 mM KCl, 5 mM $MgCl_2$) was found to perform better than NaCl and optimal values were obtained after incubation for 1 hour at 37ºC. The optimal protein concentration was determined to be 20 µM. $Rbg1_{fl}$-$Tma46_{205-345}$ and $Rbg2_{fl}$-$Gir2_{174-238}$ were analysed for the magnesium dependence with 0, 5 and 10 mM $MgCl_2$. It was seen that while there was little activity when there was no Mg^{2+} present in the reaction, comparative rates were observed at 5 mM and 10 mM Mg ion concentration.

Table 4.6: Kinetic parameters for GTP hydrolysis.

The statistics for the hydrolytic activity of the $Rbg1_{fl}$-$Tma46_{205-345}$, $Rbg2_{fl}$-$Gir2_{174-238}$, $Drg1_{fl}$-$Lerepo4_{220-426}$ complexes and the $Rbg1_{fl}$-$Tma46_{205-345}$ mutant proteins are given. The V_{max}, K_m and K_{cat} values for the deletion mutants of Rbg1 domains and helix/residues of Tma46 in the context of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex are also shown.

Protein	V_{max} *(nmol min^{-1} mg^{-1})*	K_m *(μM)*	k_{cat} *(min^{-1})*
$Rbg1_{fl}$-$Tma46_{205-345}$	0.0593 ± 0.0018	304.1 ± 32.5	0.0034
$Rbg2_{fl}$-$Gir2_{174-238}$	0.0514 ± 0.0055	303.3 ± 113.5	0.0026
$Drg1_{fl}$ + $Lerepo4_{220-426}$	0.1154 ± 0.0037	319.5 ± 35.0	0.0074
Rbg1-Tma46 Mutants			
$Rbg1_{1-294}$-$Tma46_{205-345}$ (ΔTGS)	0.0166 ± 0.0022	426.3 ± 165	0.0008
$Rbg1_{\Delta175-243+G}$-$Tma46_{205-345}$ (ΔS5D2L)	0.0157 ± 0.0028	221.5 ± 159.2	0.0006
$Rbg1_{fl}$-$Tma46_{239-345}$ (Tma46Δα1)	0.0370 ± 0.0012	191.2 ± 26.6	0.0020
$Rbg1_{fl}$-$Tma46_{205-345\ I241A,\ F246A}$	0.0288 ± 0.0021	428.9 ± 94.7	0.0017
$Rbg1_{fl}$-$Tma46_{205-345\ I241A}$	0.0406 ± 0.0018	186.0 ± 34.5	0.0023
$Rbg1_{fl}$-$Tma46_{205-345\ W249A\ K250E}$	0.0345 ± 0.0014	425.8 ± 52.6	0.0020

The $Rbg1_{fl}$-$Tma46_{205-345}$, $Rbg2_{fl}$-$Gir2_{174-238}$ and $Drg1_{fl}$ + $Lerepo4_{220-426}$ complexes showed weak GTP hydrolytic activity (K_{cat} of 0.0034 min^{-1}, 0.0026 min^{-1} and 0.0074 min^{-1} respectively, Figure 4.14.C, Table 4.6). The catalytic site point mutant $Rbg1_{fl\ GFPSVGKN}$ − $Tma46_{205-345}$ did not show significant hydrolytic activity above the detection limit of the assay confirming the specificity of the reaction (Figure 4.14.C). The recombinant human homologous complex, $Drg1_{fl}$-$Lerepo4_{220-396}$ was slightly more active than the cognate yeast complex.

4.7 Effect of the DRG-DFRP structural components on GTPase activity

4.7.1 DFRP interaction and GTPase activity

In comparison to the complex proteins, Rbg1 and Drg1 showed significantly reduced GTPase activity on their own (Figure 4.14.C). The increased activity of the complex over free Rbg1 could be attributed mostly to the fact that Rbg1 might be a more stable protein in presence of Tma46 (seen also by the increased Tm for the complex in thermofluor). This is evidenced by a decrease in activity of Rbg1 in complex with a $Tma46_{205-345}$ mutant lacking the helix α1 (Table 4.6). Furthermore, mutants of residues of

Tma46 in the interface between Rbg1 (helix α8) and Tma46 (helix α2) such as Ile241, Phe246, Trp249 and Lys250, the latter three of which are involved in pi stacking interactions, also show reduced activity. Curiously, Drg1 also showed reduced activity alone, but incubating the same Drg1 protein with Lerepo4$_{220\text{-}396}$ for 10 min prior to the assay increased the GTP hydrolytic rates. Similar catalytic values were also obtained with Drg1$_{fl}$-Lerepo4$_{220\text{-}396}$ co-purified complex.

4.7.2 Effect of DRG domain deletions on GTPase activity

A deletion mutant of Rbg1 lacking the TGS domain in complex with Tma46-dfrp (Rbg1$_{1\text{-}294}$-Tma46$_{205\text{-}345}$) was found to show GTP hydrolytic activity albeit low suggesting that TGS domain might not be entirely necessary for GTPase activity although they might be affecting the stability of the complex. The statistics are as given in table 4.6. Rbg1$_{\Delta 175\text{-}243+G}$-Tma46$_{205\text{-}345}$ lacking the S5D2L domain were also analysed for the effect of the S5D2L domain on GTPase activity. Although reduced, the activity was seen to be not completely abolished.

Since the hydrolytic activity of the DRG proteins was very low compared to the classic GTPases and happened in the absence of any activating proteins, we wanted to analyse if RNA binding would stimulate hydrolytic activity *in vitro*. Accordingly, the assay was done in the presence of RNA by incubating Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ with poly(U) for 20 min prior to the addition of GTP. However, no increase in the catalytic activity was observed. The kinetic parameters were similar to the native protein suggesting that non-specific RNA homopolymer binding might not influence Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ GTPase activity *in vitro*.

4.8 Structure of the GTPase inactive mutant and guanine nucleotide bound complex

The structure of the inactive protein containing mutation on the G1 motif (GFPSVGKS → GFPSVGK**N**) published previously to be essential for function (Daugeron *et al.,* 2011) and shown in section 4.6.2 to be devoid of catalytic activity was solved to 2.6 Å resolution. The structure of the mutant showed the mutated residue in the active site pocket probably able to accommodate GTP (also evidenced by the thermal shift assay result) although incapable of catalysing the hydrolytic reaction. The Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ complex structures bound to nucleotides GDP and GTPγS were also solved to 2.9 and 2.5 Å resolutions respectively.

Given that the cell parameters for the Rbg1$_{fl\ GFPSVGK\mathbf{N}}$-Tma46$_{205\text{-}345}$ and Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ GDP/ GTPγS data were similar to the Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ native crystals, a rigid body refinement was enough to obtain initial electron density maps. The structures were refined using the same TLS groups defined for the native heterodimer described in

section 4.3. The statistics for the data collection and after the refinement, for the three structures are given in table 4.7.

Table 4.7: Data collection and refinement statistics of $Rbg1_{fl\ GFPSVGKN}$-$Tma46_{205-345}$ and $Rbg1_{fl}$-$Tma46_{205-345}$-GDP/GTPγS complexes.

Data Collection	Inactive mutant	GDP cocrystal	GTPγS cocrystal
No of crystals used	1	1	1
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell parameters			
a, b, c (Å)	86.09, 225.57, 85.16	85.98, 224.91, 84.93	86.29, 225.05, 84.83
α, β, λ (º)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength (Å)	0.9763	1.007	0.9763
Completeness (%)	99.74	99.8	99.5
Mean $I/\sigma(I)$ [a]	18.7	14.7	27.1
R_{meas} (%) [b]	5.3	13.2	3.0
Refinement			
Resolution (Å)	85.16 – 2.6	84.93 – 2.87	112.53 – 2.3
No. of reflections R_{work}	49105	36478	70071
Reflections used in R_{free}	2632	1923	3720
R_{work} [c]	19.1%	18.09%	21.36%
R_{free} [d]	22.8%	21.72%	23.43%
Mean B-factor-Overall	41.088	41.684	38.391
Rmsd [e]			
Bond lengths (Å)	0.0075	0.0080	0.0068
Bond angles (º)	1.2390	1.3071	1.1719

[a] Mean $[I/\sigma(I)]$ is the average of the relation between the intensity of the diffraction and the background.

[b] $R_{meas} = \{\Sigma_{hkl}\,[N/(N-1)]^{1/2}\,\Sigma_i\,|I_i(hkl) - \langle I(hkl)\rangle|\}\ /\ \Sigma_{hkl}\,\Sigma_i\,I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl)\rangle$ are the average intensities and N is the multiplicity of reflection *hkl*.

[c] $R_{work} = \Sigma_{hkl}\,\{[F_{obs}(hkl)] - [F_{calc}(hkl)]\}\ /\ \Sigma_{hkl}\,[F_{obs}(hkl)]$, where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.

[d] R_{free} corresponds to R_{work} calculated using 2.5% of the total reflections selected randomly and excluded during refinement.

[e] *Rmsd* is the root mean square deviation.

For the datasets corresponding to the GDP/GTPγS bound structures, the Fo-Fc maps before placing the nucleotides suggested electron densities in the binding pocket of Rbg1 as given in figure 4.15.A. Accordingly, for the GDP-bound data, refinements were carried out including a GDP template in the Rbg1 molecule in the open conformation of the $Rbg1_{fl}$-$Tma46_{205-345}$ structure. The ppGpp molecule from the *B. subtilis* Obg structure (PDB id: 1LNZ) after superimposition of the protein to Rbg1 and removing the two phosphates attached to the ribose sugar was taken as the template GDP molecule for the refinement. Likewise, the GTPγS molecule from transducin-alpha (PDB id: 1TND) was used for the

$Rbg1_{fl}$-$Tma46_{205-345}$-GTPγS dataset. Although the electron densities for the α and β phosphates of the guanine nucleotide were clearly seen for both GDP and GTPγS bound structures, sufficient electron density was not seen for the γ phosphate of GTPγS.

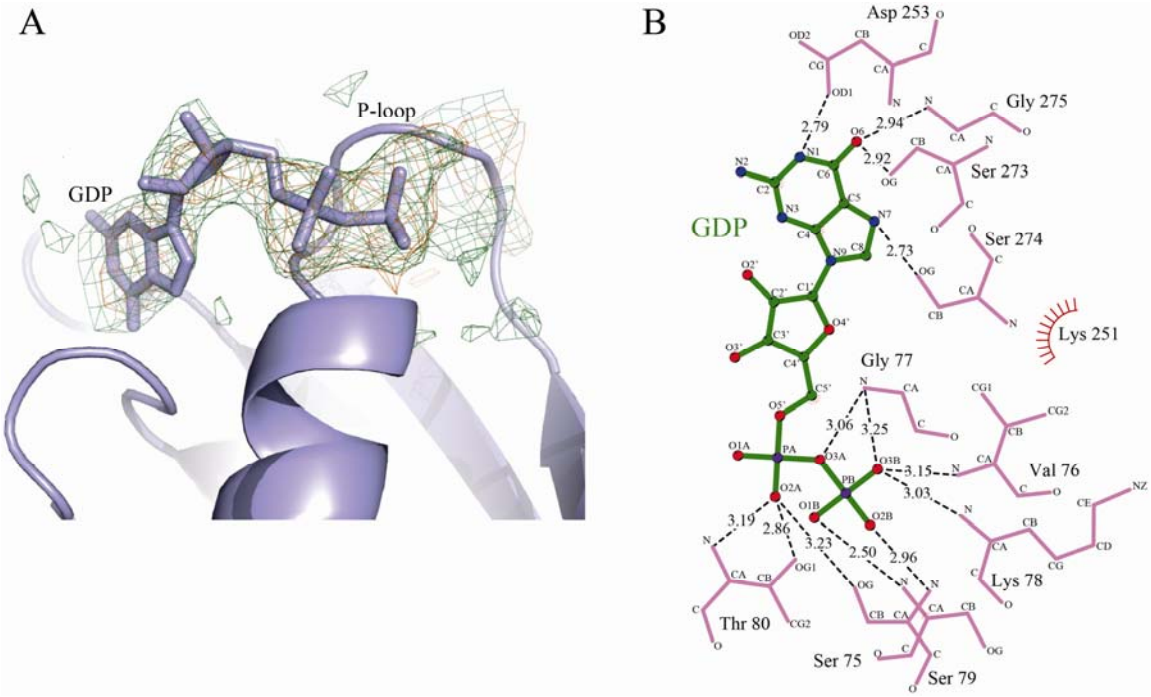


Figure 4.15: GDP/GTPγS bound $Rbg1_{fl}$-$Tma46_{205-345}$.

A. The difference electron density maps (Fo-Fc) of the GDP/GTPγS bound structures contoured at 2.5 σ are shown in orange/green respectively. The $Rbg1_{fl}$-$Tma46_{205-345}$ structure later on refined using the GDP molecule from BsObg structure is given showing the GDP in the active site for clarity.

B. Schematic representation of the contacts between GDP (green) and Rbg1 (pink) in the final refined $Rbg1_{fl}$-$Tma46_{205-345}$-GDP complex structure is given. The figure was generated using Ligplot with the hydrogen bonds shown as black dotted lines labelled with the distances given in Å and the hydrophobic interactions in red lines.

4.9 Binding of DRG-DFRP proteins to polyuridylic acid

We next analysed the capability of the recombinant DRG and DFRP proteins to bind RNA. For this purpose, RNA homopolymer binding assay as previously described for *Xenopus laevis* Drg1 and Drg2 (Ishikawa *et al.,* 2003) was performed. Indeed, Rbg1/Rbg2 in complex with Dfrp domain of Tma46/Gir2, co-elutes with polyuridylic acid (poly(U)) bound agarose beads. As negative control, BSA was used which did not show any binding. Beads treated with RNAse A prior to addition of the protein completely eliminated binding confirming that there is no non-specific binding to the agarose beads (Figure 4.16.A).

The binding also took place in a competitive manner in the presence of free poly(U). As shown in the figure, with increasing concentrations of free poly(U) added to the protein before binding to the beads, a decrease in binding was observed as probably the RNA binding site on the protein gets saturated. Also, addition of increasing concentrations of heparin which is an RNA competitor and NaCl disrupts the binding only somewhat, indicating that the protein has strong affinity for poly(U) (Figure 4.16.B). On an added note, the GTPase inactive mutant $Rbg1_{fl\ GFPSVGK\mathbf{N}}$-$Tma46_{205-345}$ also bound to poly(U) confirming that RNA binding could occur independent of the presence of GTP or hydrolytic activity.

Binding to RNA by the DRG proteins also did not seem to require the presence of DFRP as $Rbg1_{fl}$ and $Drg1_{fl}$ alone could bind to the poly(U) beads. Among the DFRP proteins, $Tma46_{fl}$ on its own could not be tested for binding since it degrades highly when expressed alone. But Lerepo4 (human Dfrp1 counterpart), although the full length could not be expressed, the N- (1-220) and C-terminals (220-396) were tested. While $Lerepo4_{220-396}$ which is the C-terminal does not show any binding, the N-terminal does bind. This could be due to the presence of the Zinc fingers in the N-terminal of the Dfrp1 proteins (Lerepo4 and Tma46). However, $Gir2_{fl}$ (yeast Dfrp2) does not show any binding to poly(U) beads. This result is supported by the fact that $Gir2_{fl}$ does not contain any Zinc fingers at the N-terminal, having an RWD domain instead.

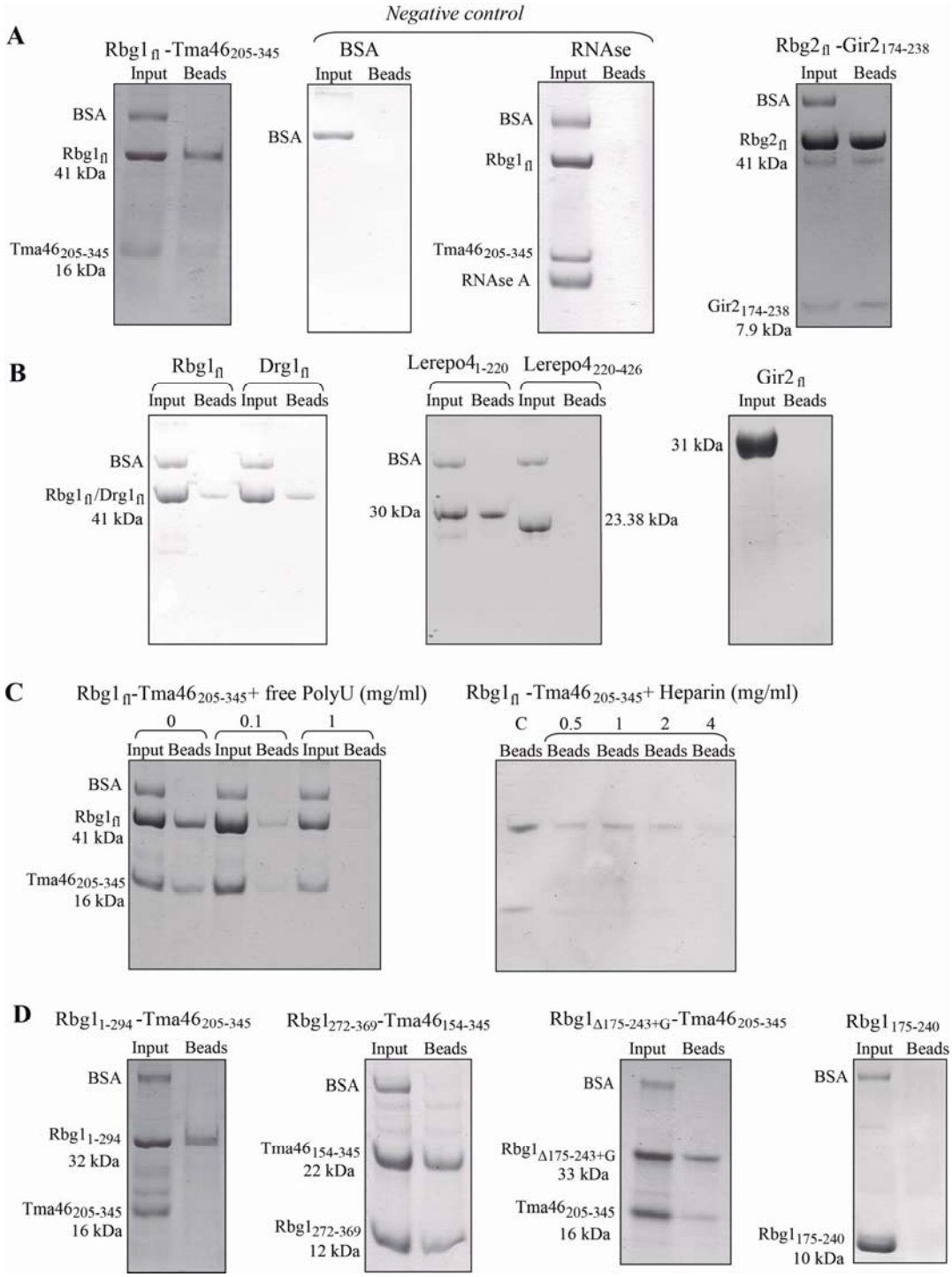


Figure 4.16: Non-specific RNA homopolymer binding of the DRG-DFRP proteins.

A. Drg-Dfrp complexes bind to RNA: $Rbg1_{fl}$-$Tma46_{205-345}$ and $Rbg2_{fl}$-$Gir2_{174-238}$ complex pulled down with poly(U) beads (lane Beads). Binding of Rbg1-$Tma46_{205-345}$ to RNase treated beads and BSA are included as negative controls.

B. Binding can be saturated and is stable: Increasing concentrations (0.1, 1 mg/ml) of free poly(U) added to Rbg1-Tma46$_{205-345}$ before binding to the poly(U) agarose beads competed for binding demonstrating that binding occurs through a defined surface that can be saturated. Addition of increasing concentrations of heparin (0.5, 1, 2 and 4 mg/ml) to the washes after Rbg1-Tma46$_{205-345}$ was bound to the beads did not completely displace the protein from the beads indicating that binding is stable.

C. Rbg1$_{fl}$ and Drg1$_{fl}$ (DRGs) show binding to poly(U). Gir2$_{fl}$ and Lerepo4$_{220-396}$ does not show binding whereas Lerepo4$_{1-220}$ (N-terminal) can be seen to be eluted with the poly(U) beads.

D. Contribution of individual domains of Rbg1: Rbg1$_{1-294}$ (ΔTGS) and Rbg1$_{272-369}$ (TGS), both in complex with Tma46 (dfrp) fragment and Drg1$_{289-367}$ (TGS) showed binding to poly(U) beads suggesting that TGS domain can bind to poly(U) but is not essential for Rbg1 RNA binding. Rbg1$_{\Delta 175-243+G}$ –Tma46$_{205-345}$ (ΔS5D2L) bound to poly(U) beads while Rbg1$_{175-240}$ (S5D2L) domain alone did not show any binding.

The S5D2L domain alone of Rbg1 andRbg2 and Drg1 did not bind to poly(U) whereas the complex lacking the S5D2L showed binding. Rbg1-Tma46 complex lacking the TGS domain (Rbg1$_{1-294}$-Tma46$_{205-345}$) was seen bound to the beads, although Tma46$_{205-345}$- was not seen in the eluate. This might be probably due to the highly acidic nature of Gir2 (also contributing to its anomalous electrophoretic behaviour in a gel) causing an electrostatic repulsion with negatively charged RNA added onto the fact that binding to DRGs is weakened by the deletion of the TGS domain. Additionally, TGS domain on its own was found to be able to bind to poly(U) as evidenced by Drg1$_{289-367}$ (TGS). Rbg1 TGS domain (Rbg1$_{272-369}$) in complex with the longer Tma46 construct (Tma46$_{154-345}$) bound to poly(U). The latter binding however cannot be attributed to be due to either Rbg1$_{272-369}$ or Tma46$_{154-345}$ alone.

4.10 Rbg1-Gir2 functional complementation in yeast

Rbg1-Tma46 and Rbg2-Gir2 are have been reported to be functionally independent complexes having a redundant role in a biological process along with the RNA helicase Slh1 as described in detail in section 1.6.1. Even so, there are some evidences which suggests that Rbg1 might also be able to interact with Gir2 apart from its binding partner, Tma46 (Wout *et al.,* 2009). In our hands, co-purification of Rbg1$_{fl}$ and Gir2$_{fl}$ proteins expressed from co-transformed plasmids ($_{6His}$Rbg1$_{fl}$ (KanR) and Gir2$_{fl}$ (AmpR)) in *E.coli* showed that the proteins form a complex *in vitro*. The chromatograms of the purification procedure involving first the nickel chelating affinity chromatography followed by size exclusion are provided in figure 4.17. The elution volume suggests that the Rbg1-Gir2 complex eluted as a single heterodimer.

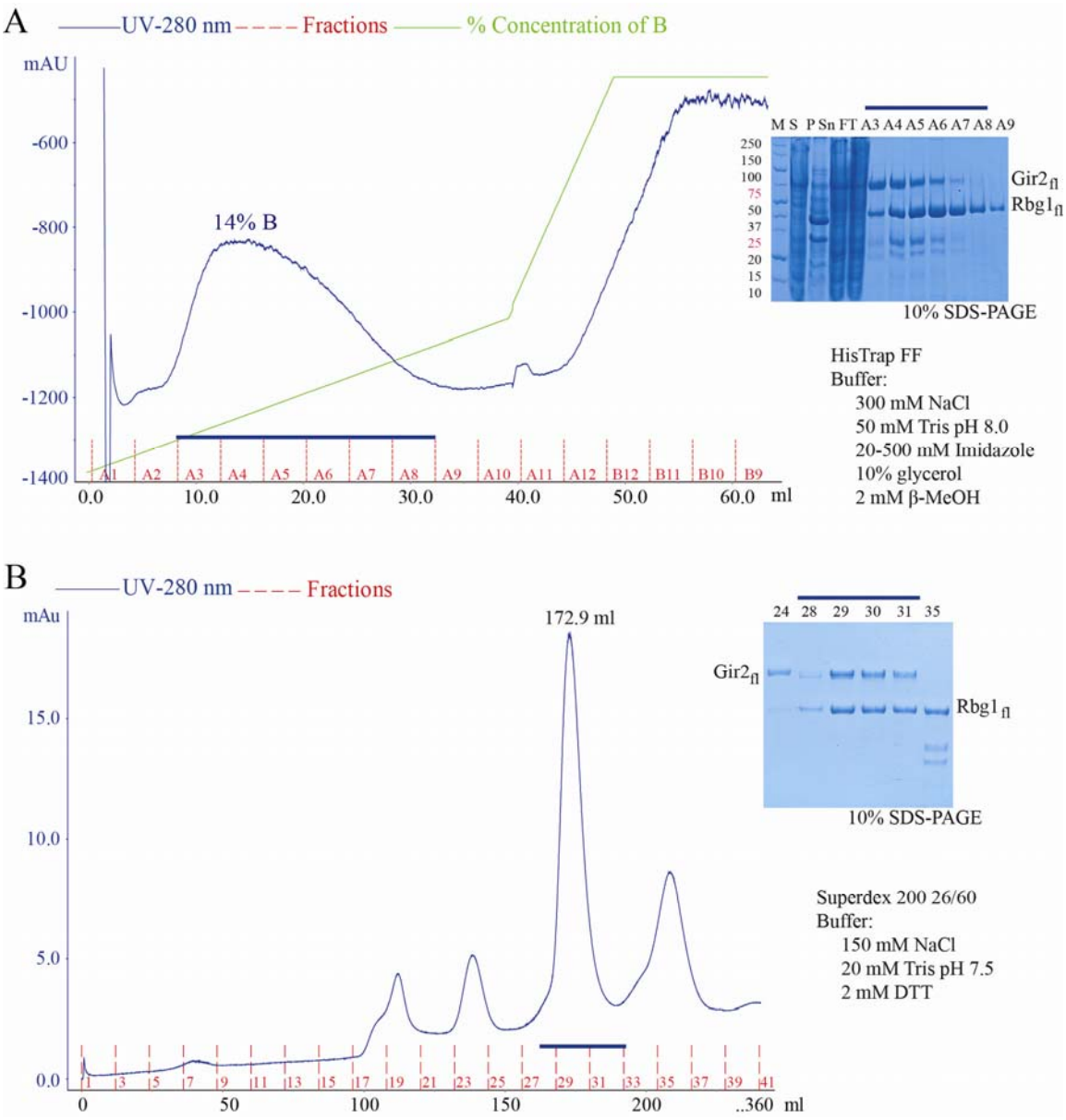


Figure 4.17: Purification results for $Rbg1_{fl}$-$Gir2_{fl}$.

A. The co-expressed $Rbg1_{fl}$ protein containing the 6xHis tag bound to the HisTrap column and the $Gir2_{fl}$ protein could be eluted along with $Rbg1_{fl}$ even though it lacked the 6xHis tag.

B. The $Rbg1_{fl}$-$Gir2_{fl}$ complex (fractions 28-31) could be separated from the individual $Rbg1_{fl}$ (fraction 35) and $Gir2_{fl}$ (fraction 24) proteins in the size exclusion chromatography step.

A study on budding yeast, *Saccharomyces cerevisiae* was undertaken in order to study whether this interaction takes place *in vivo*. Daugeron *et al.* have previously demonstrated that the triple deletion mutant yeast strain *Δrbg1Δrbg2Δslh1* and *Δtma46Δgir2Δslh1* grow slowly compared to a wild type strain (Daugeron *et al.,* 2011). The experimental question taken up here was whether a yeast strain lacking *rbg1*, *rbg2*,

tma46 and *slh1* from its genome and containing only *gir2*, upon complementation with wild type *rbg1* would recover normal growth. For this purpose, the quadruple mutant strain *Δrbg1Δrbg2Δtma46Δslh1* was constructed by crossing the two haploid yeast strains *Δrbg1Δtma46Δslh1* and *Δrbg2Δslh1* obtained from Dr. Bertrand Séraphin. Sporulating the resulting diploids gave haploids which were analysed on selective plates to identify the quadruple mutant strain (Refer section 3.5).

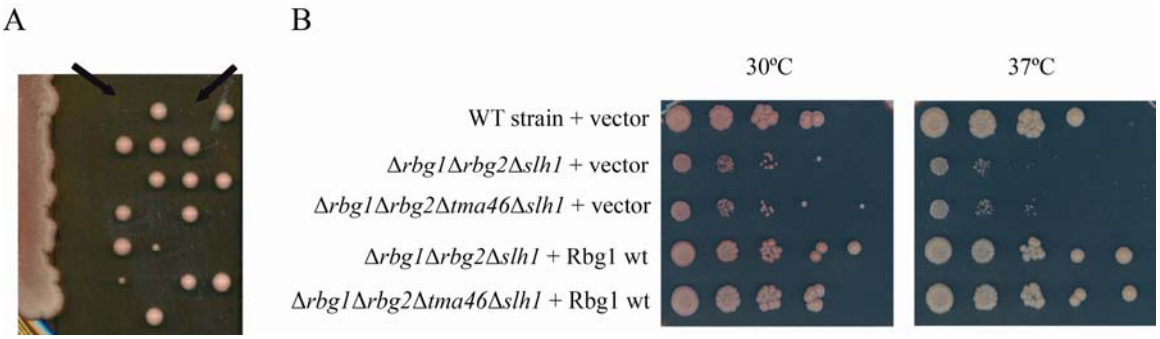


Figure 4.18: Tetrad dissection and yeast growth phenotype assay.

A. Tetrads were dissected in order to obtain the quadruple strain. Proper segregation of the four gene deletions (*Δrbg1*, *Δrbg2*, *Δtma46* and *Δslh1*) in the tetrads were identified using the marker genes (*HISMX6*, *KANMX4*, *HISMX6* and *URA3Kl* respectively) inserted in these positions. Around 50 tetrads were dissected in order to obtain a 2:0 segregation which assures the presence of both *Δrbg1* and *Δtma46* deletions as they had the same marker.

B. Comparison of growth phenotype of the triple and quadruple mutant strains complemented with empty vector or the Rbg1 wild type protein. The wild type strain transformed with the empty vector is given as the positive control. The plates were photographed after incubation for 4 days at 30 and 37ºC.

Drop assay was carried out in order to compare the growth phenotype of the mutant and complemented strains. As can be seen in figure 4.18, indeed the quadruple *Δrbg1Δrbg2Δtma46Δslh1* mutant showed similar negative growth phenotype as that of the triple *Δrbg1Δrbg2Δslh1* mutant in comparison to the wild type as expected. Both the above triple and quadruple mutant when transformed with a plasmid expressing wild type Rbg1 (YCplac111-*HA-RBG1* (*ARS-CEN LEU2*)) restored normal growth which was comparable to that of the wild type strain (BSY1664, table 3.5) transformed with a control plasmid (YCplac111).

Whereas Rbg1 complementation of the triple *Δrbg1Δrbg2Δslh1* mutant strain signifies restoration of growth due to the functional Rbg1-Tma46 complex formation, this is not the case with the quadruple mutant as Tma46 is absent in this strain. This suggests that the recovery of normal growth rate is most probably due to a functional complex formation between Rbg1 and Gir2.

4.11 Functional study by translation reporter assays

Although evidences suggest an involvement of the DRG proteins in translation, nothing is known yet about its exact role. In order to obtain functional clues, preliminary assays measuring efficiency of termination and fidelity of elongation which are parts of the protein synthesis process were done using luciferase reporter genes in triple mutant Δ*rbg1*Δ*rbg2*Δ*slh1* and wild type (BSY1664) yeast strains *in vivo*. The assay was carried out using the dual luciferase reporter assay system which measures the ratio between firefly and *Renilla* luciferase activity (described in section 3.5.6). Plasmids expressing the *Renilla* and firefly luciferase genes under the control of the *S. cerevisiae* PGK promoter (Salas-Marco and Bedwell 2005) were transformed into the wild type and Δ*rbg1*Δ*rbg2*Δ*slh1* yeast strains.

4.11.1 Translation fidelity using the luciferase assay

For assaying the frequency of misreading in the elongation step of protein translation *in vivo*, the dual luciferase system in which the firefly luciferase gene had an inactivating mutation was used (Kramer and Farabaugh 2007, Kramer *et al.,* 2010). An essential, highly conserved lysine (K529) in the active site of firefly luciferase was mutated to any one of the following amino acids: Phe (UUU), Stop (UAA/UAG/UGA), Gln (CAA/CAG), Glu (GAG), Ile (AUA), Met (AUG), Thr (ACA/ACG), Arg (AGA/AGG), Asn (AAU/AAC). Firefly luciferase activity can be detected depending on the frequency of every near-cognate misreading error by , i.e. misincorporation of the wild type lysine residue at the mutated codon restores firefly luciferase activity.

Yeast strain expressing *Renilla* and firefly luciferases from a plasmid containing codon AAA (Lys) at amino acid position 529 of the firefly protein was used as the positive control. Codon UUU which codes for phenylalanine was used as the negative control as it is incapable of making any Watson-Crick base pairs with the UUU anticodon of $tRNA^{Lys}$ and has the minimum probability of reading the codon as lysine. The frequency of misincorporation of this lysine is measured as the luciferase activity. The standardized Fluc activity (Fluc/Rluc) of the mutant firefly luciferase enzyme with respect to the wild type luciferase was used to quantify the translational fidelity rates in the strain expressing each mutant firefly luciferase. Unfortunately, no significant difference in the expression of the mutated luciferases between the mutant and wild type strains (figure 4.19.B) was observed.

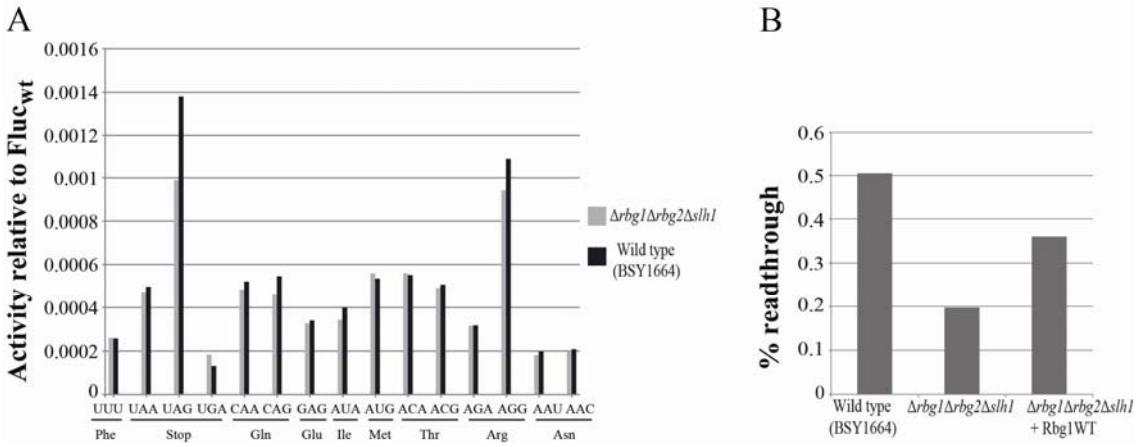


Figure 4.19: Translational fidelity and stop codon recognition reporter assays.

A. Fidelity assay results measuring efficiency in proof reading in the translation process. Given on the y axis is the firefly luciferase K529 mutant activity relative to the wild type Fluc and on the x axis are the Δ*rbg1*Δ*rbg2*Δ*slh1* (given in grey) and wild type (black) strain samples containing the mutated codons which replace for the lysine 529 of firefly luciferase.

B. Reporter assay measuring the efficiency of stop codon recognition. The percentage readthrough of the wild type, Δ*rbg1*Δ*rbg2*Δ*slh1* triple mutant strain and the Δ*rbg1*Δ*rbg2*Δ*slh1* strain complemented with the wild type Rbg1protein is given.

4.11.2 Assay for stop codon recognition

One of the expression plasmids transformed contained a stop codon (UAA) between the *Renilla* and firefly luciferase genes and the other contained a glutamine (CAA) instead of the stop codon. In a wild type yeast strain expressing the two luciferases with the stop codon (UAA) in between, the ratio of firefly to *Renilla* luminescence (Fluc/Rluc) measured would be minimal, as proper stop codon recognition takes place and firefly gene is not translated. Whereas in the wild type strain expressing the luciferases with the glutamine (CAA) mutation, firefly protein is translated along with *Renilla* luciferase as a fusion product which thereby increases the Fluc/Rluc ratio. The percentage of readthrough is calculated as Fluc/Rluc (UAA→Stop) divided by Fluc/Rluc (CAA→Glu) X 100). Improper stop codon recognition at UAA expresses the firefly luciferase resulting in an increase in the percentage of readthrough. Comparison of the readthrough of the wild type and triple deletion (Δ*rbg1*Δ*rbg2*Δ*slh1*) mutant strain would give an idea about the involvement of the Rbg1/Rbg2/Slh1 proteins in translation through their effect on stop codon recognition.

The assay demonstrated that the triple mutant strain in fact causes a decrease in the percentage of readthrough of the luciferase protein as compared to the wild type strain as can be seen in figure 4.19.A. Although this decrease is modest, being only 2-3 folds, it is interesting to note that on complementation with the wild type Rbg1 protein the percentage read-through tended to increase.

Discussion

5. Discussion

An understanding of the molecular and cellular function of the innumerable number of proteins in an organism is critical for combating many diseases. Unlike DNA, the three dimensional structure is unique to each protein and paves the way to its function. Structural information of a protein can provide insights into the biophysical and biochemical roles, mechanism of action and how it networks within the cell which ultimately helps in designing new drugs for various disease conditions. Biological macromolecular structure solution still relies largely on X-ray crystallography as the most widest and precisely used method of choice. The major bottlenecks of this method, however, is the procurement of diffraction quality crystals and another being the "phase problem" in case of proteins with no homologous structures available.

5.1 Analysis of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex crystal forms and packing

In the present study, we have solved the X-ray crystal structure of the yeast Rbg1 protein in complex with the C-terminal part of Tma46. Analyses of the crystal packing in the unit cell showed that each of the domains interacts with itself from the corresponding symmetry related molecule. Accordingly, the G-domain of molecule A interacts with G-domain of molecule A´ (symmetry related molecule) while that of molecule B interacts with the symmetry related molecule B´ (Figure 5.1.A). Other interfaces in the crystal lattice are also curiously formed between the same domains in adjacent molecules, i.e. the HTH-S5D2L and the TGS domain from symmetry related A and B molecules lie next to each other. Similarly Tma46 interacts with its symmetry related molecule through the α2 helix.

The crystals of $Rbg1_{fl}$-$Tma46_{205-345}$ seemed to have a flat interface between the symmetry related molecules looking down the c axis (G- and TGS domain interface) which could lead to sliding within the lattice. A resolution dependent sharp drop in I/σI together with relatively high mosaicity values were seen for these crystals which might be explained by the presence of this flat interface.

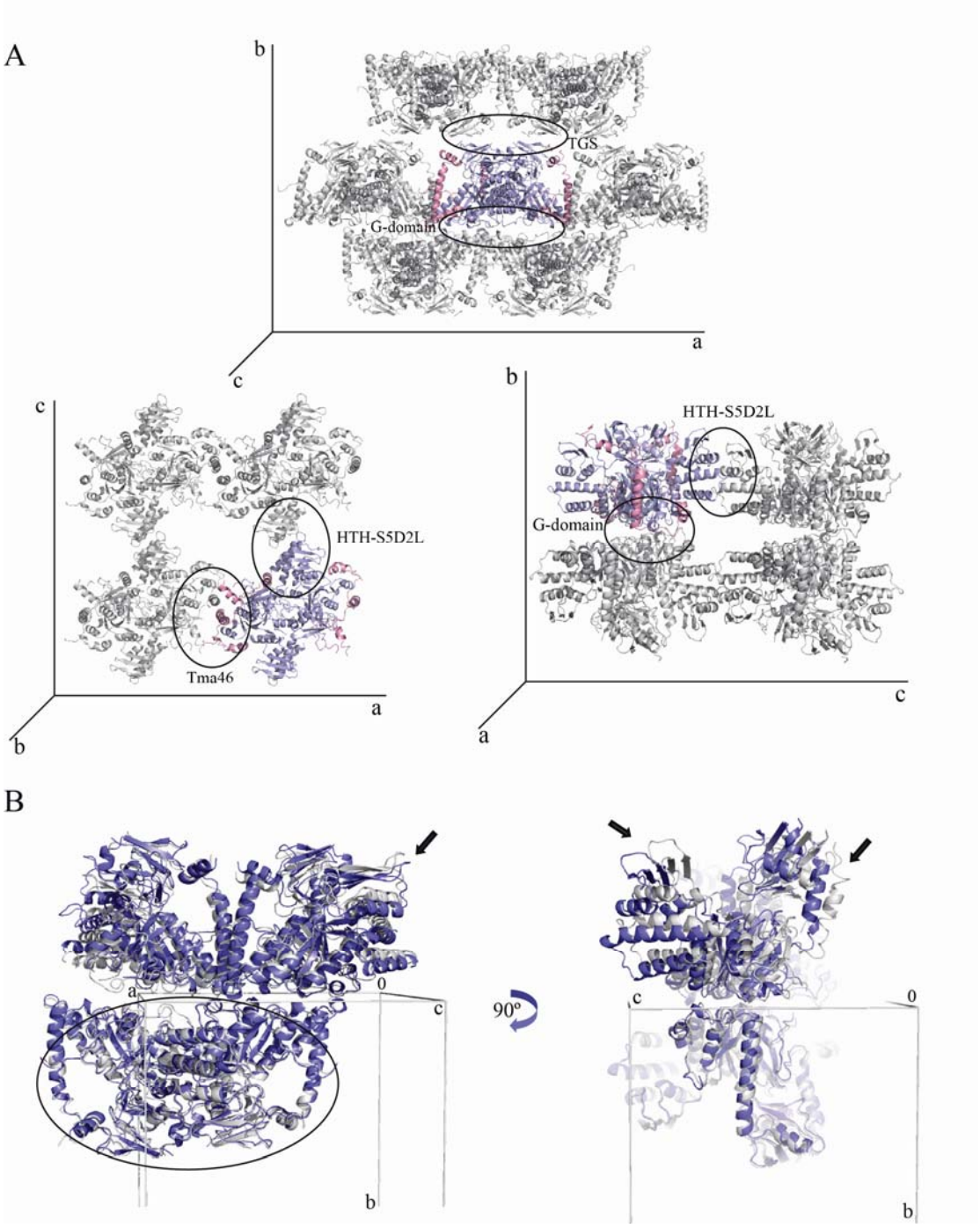


Figure 5.1: Packing in the Rbg1$_{fl}$-Tma46$_{205-345}$ crystals.

A. The arrangement of the symmetrical molecules showing the interacting surface between the G- and TGS domain down the c axis, the HTH-S5D2L and Tma46 down b axis and between the HTH-S5D2L and G-domain looking down the a axis.

B. Comparison of the packing in the monoclinic (grey) and orthorhombic (blue) crystal lattices is shown. The asymmetric unit of the $P2_12_12$ crystal structure was aligned (SSM superpose) to the corresponding two heterodimers of the $P2_1$ structure (given inside the circle). The displacement between the two lattice structures is indicated by the arrows.

Curiously, $Rbg1_{fl}$-$Tma46_{205-345}$ crystallized in two closely related unit cell dimensions in our hands belonging to primitive monoclinic and orthorhombic lattices. The occurrence of these two lattices was observed in similar crystallization conditions, monoclinic crystals obtained in 2.8 M Sodium citrate at 21ºC and the orthorhombic crystals obtained in 2.38 M Sodium formate, 0.21-0.25 M Sodium citrate at 4ºC. This change in crystal lattice was also earlier observed in crystal dehydration experiments where the unit cell rearranged from near-orthorhombic to orthorhombic. The diffraction pattern of the $Rbg1_{fl}$-$Tma46_{205-345}$ crystals belonging to the monoclinic lattice suggested higher internal disorder and mosaicity compared to the primitive orthorhombic. The solvent content of these crystals was high (around 60%) which would have facilitated the transition from the monoclinic lattice with four molecules in the asymmetric unit to the more ordered orthorhombic lattice in the dehydration experiments. Comparatively slower crystal growth at 4ºC would have facilitated the protein molecules to arrange in a more ordered fashion providing lesser mosaicity and a higher symmetry crystal lattice. Despite that both dehydration and slow crystal growth provide a primitive orthorhombic lattice, they are not identical since data from the slow growing crystals processed correctly as orthorhombic unlike the dehydration dataset. With the addition of Sodium citrate to the reservoir condition of the monoclinic crystals, there were some occurrences of a heterogeneous population of monoclinic/orthorhombic crystals even at 21ºC. A combination of the presence of Sodium citrate as an additive and slower crystal growth at 4ºC allowed the obtainment of the primitive orthorhombic crystals that in the long run facilitated us to solve the structure.

The arrangement of $Rbg1_{fl}$-$Tma46_{205-345}$ molecules in the monoclinic crystal was analysed respective to the orthorhombic cell. For this purpose, molecular replacement (Phaser) using the heterodimer in the $P2_12_12$ (orthorhombic) structure as a search model was performed which fit four heterodimers in the asymmetric unit. This was followed by a few refinement cycles (data not shown). The primitive monoclinic crystals were found to belong to the $P2_1$ space group. Despite the fact that there is no obvious difference in the crystal contacts between the molecules in the two spacegroups, a comparison of the packing in the $P2_1$ and $P2_12_12$ crystals showed that there is a slight progressive displacement of the adjacent molecules along the shortest axis (a axis) in the crystal belonging to the $P2_1$ spacegroup; corresponding is the c axis in $P2_12_12$ (Figure 5.1.B). This breaks the perpendicularity between this axis and the longest axis (as in an orthorhombic cell) thus contributing to the monoclinic crystal lattice.

5.1.1 The GTP binding site shows flexibility

In GTP-binding proteins, G2 (Switch I) and G3 (Switch II) are motifs well-known to be undergoing substantial conformational changes upon GTP/GDP binding thus transducing signals to various downstream effector proteins. The loops containing the

G2/G3 motifs in Rbg1 are probably for this reason not seen clearly in the electron density due to their high flexibility. Even when GDP or GTPγS is bound to Rbg1, the loops could not be clearly defined. This is also true in most nucleotide bound G domain structures, for e.g. the YchF hOLA1 structure bound to AMPPCP (PDB id: 2OHF) where the loop containing the G2 and G3 motifs are not observed. We believe that the $Rbg1_{fl}$-$Tma46_{205-345}$ protein complex that was crystallized either in nucleotide-bound or unbound state probably belongs to the same inactive conformation in the G-domain. It could be possible that the conformation of the active form may not be compatible with the packing occurring in the crystals obtained.

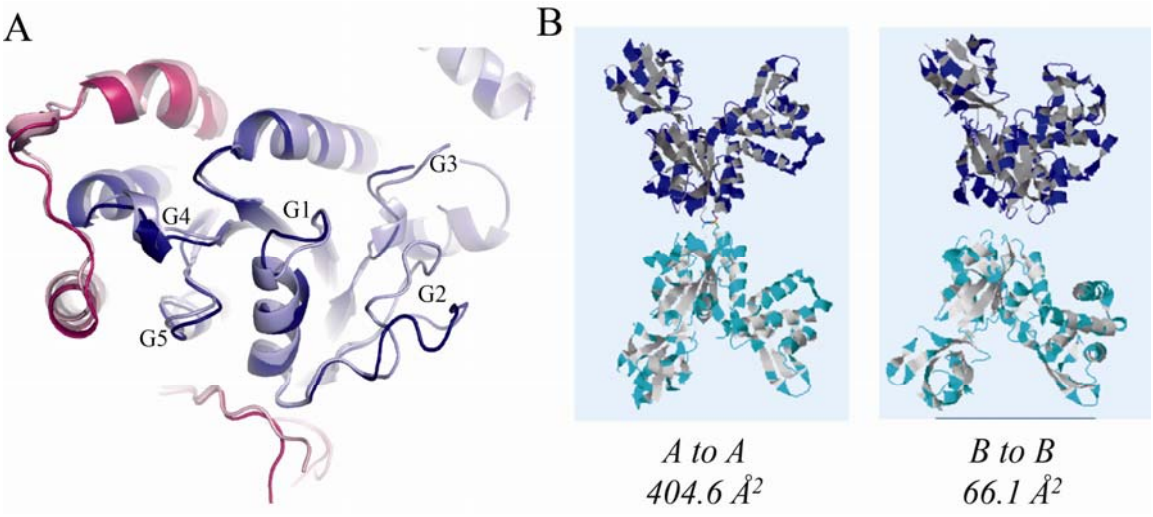


Figure 5.2: Detail of the G-domain of the $P2_12_12$ crystals.

A. Molecules A and B are superposed to show the open/closed conformation with molecule A given in light blue and molecule B in dark blue; Tma46 from both molecules is coloured in shades of pink. The five G-motifs are indicated with the loop positions different in the two molecules in G1, G2 and G3.

B. The contacts between molecule A to A′ of the symmetry related molecule and that of B to B′ is represented along with the interface areas given by Pisa.

Of the two molecules of Rbg1 in the asymmetric unit of the nucleotide free Rbg1-Tma46 complex, molecule A has the G motifs in a closed conformation when compared to molecule B which was in an open conformation (Figure 5.2.A). In the $Rbg1_{fl}$-$Tma46_{205-345}$ GDP/GTPγS bound structures, the Rbg1 molecule in the open position contained the nucleotide while the other was vacant. Despite full crystallization screens being setup for the co-crystallization of the nucleotides with the $Rbg1_{fl}$-$Tma46_{205-345}$ complex, crystals were obtained in similar conditions and had the same two conformations (open/closed) in the asymmetric unit. Comparatively, in *Bacillus subtilis* Obg structure (PDB id: 1LNZ), which is sequentially closer to the Rbg1 G-domain, two monomers were present in the asymmetric unit, one in the nucleotide bound state and the other unbound. The nucleotide however, was not seen until the NCS restraints were released (Buglino *et al.,* 2002). The

presence of nucleotide free/bound states or two different nucleotide bound states in the same crystal have also been observed in many other cases in the literature (Rudolph *et al.,* 1999, Savvides *et al.,* 2003). The mechanistic significance of this observation was however not studied. In our case, analysis of the crystal packing showed that this difference in the Rbg1 GTP binding pocket between the two molecules in the asymmetric unit could also be due to contacts with the symmetry related molecules in this area. The GTP binding region in molecule A which makes contacts with that in molecule A´ was found to be more closer together than in molecule B to B´ (Figure 5.2.B). Pisa analysis shows that A-A´ interface area is about six times that of B-B´. Nevertheless, whether this close packing is the only reason for the open/closed conformation is a matter of speculation. The formation of the crystal, preferentially incorporating two different nucleotide binding states of the same GTPase molecule in equal proportion, indicates that the overall contribution of both states to crystallization is similar despite that the loop containing the G2 motif in the closed conformation is better defined than that in the open conformation where residues 92 to 97 are not seen in the electron density maps. In the primitive monoclinic crystal structure, of the four molecules in the asymmetric unit, two were in the open position and two in the closed as expected.

5.2 Structure based evolutionary and functional analysis of the DRG proteins indicates closer similarities to translation factors

The **D**evelopmentally **R**egulated **G**TP (DRG) binding protein subfamily belongs to the Obg family of the TRAFAC class GTPases. Globally, GTPase proteins are characterised by the presence of the G-domain and the ability to hydrolyse GTP to GDP being implicated in various cellular processes. The Rbg1 G-domain was found to be highly similar to the well conserved GTP binding domain of the P-loop GTPases, consisting of a core α/β unit with the helices surrounding the central sheet. In particular, Rbg1 was similar to proteins Obg, YchF, FeoB, HflX, Ras-related proteins or Era, etc. sharing a characteristic antiparallel beta-strand topology at the Walker B motif and functioning as monomers unlike the SIMIBI class of GTPases which function as dimers (Leipe *et al.,* 2002).

Apart from the G-domain, the TRAFAC class member proteins are usually comprised of other domains in addition. Largely, very little conservation exists in sequence, structure or function between the TRAFAC member proteins including the OBG family outside the G-domain, the latter family being structurally not very well characterised (Figure 5.3). Although unified by a common working mechanism of GTP hydrolytic activity, these additional elements helps in providing the diverse physiological functions that the TRAFAC class members can perform. Structural information from this work reveals that DRGs in this sense, have acquired multiple elements evolutionarily. It was found to be a multi-modular protein having the G-domain and three additional domains, TGS, HTH and S5D2L domains.

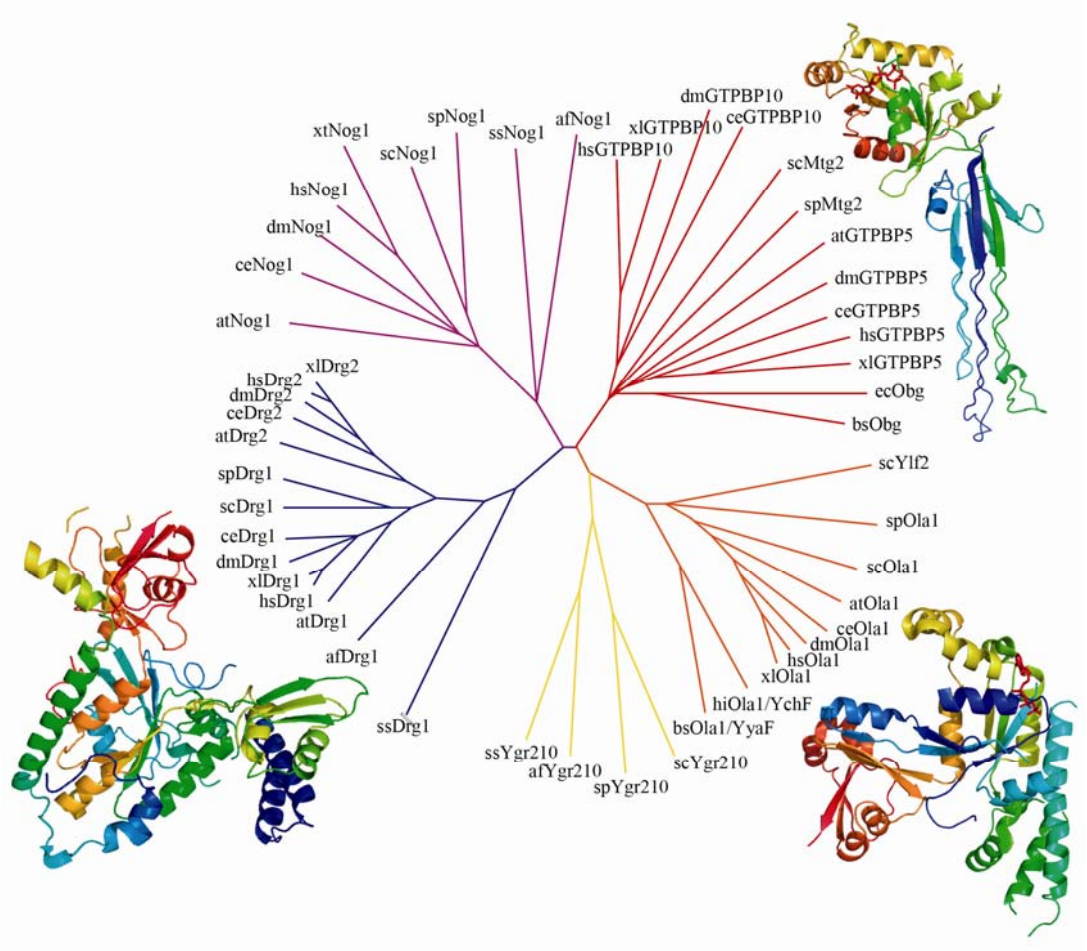


Figure 5.3: The phylogenetic tree of the OBG family. The available structures from *Bacillus subtilis* Obg protein (PDB id: 1LNZ), human YchF OLA1 (PDB id: 2OHF) and the $Rbg1_{fl}$-$Tma46_{205\text{-}345}$ structure solved in the present study (PDB id: 4A9A) is shown next to the corresponding branch representing its subfamily.

As according to the current information available, the structural similarity of the DRG protein to its fellow OBG family member proteins is restricted to the core G-domain except the YyaF/YchF subfamily, which is the only other OBG family member known so far structurally to contain the TGS domain. The YyaF/YchF and DRG subfamilies belong to each of the two major branches of the OBG family, the Ygr210-YyaF/YchF and the Obg-DRG(NOG1) and thus might be the ones most closely resembling the ancestor of the OBG family that was present in LUCA (Leipe *et al.,* 2002).

The TGS domain was named so after its presence in ThrRS, GTPase and SpoT proteins. Although found in many proteins its function is still unknown and is characterised to be related to the ubiquitin superfamily as it is seen to be structurally similar to ubiquitin. The *E. coli* ThrRS TGS domain was also previously shown to be homologous to αL motif of Hsp15 (Staker *et al.,* 2000) and the ribosomal S4 protein (Dean and Nomura 1980,

Mizushima and Nomura 1970). Rbg1 TGS domain superimposes to YchF human OLA1 (PDB id: 2OHF) TGS domain with an RMSD of 2.2 Å over main chain atoms. In hOLA1 and Rbg1 structure, the orientation of the TGS domain with respect to the G-domain is very similar; however, in place of the HTH and S5D2L domains of Rbg1, hOLA1 has an inserted coiled coil domain (Koller-Eichhorn *et al.,* 2007). An additional helix present in hOLA1 TGS is partly formed by helix 3 of Tma46 as shown in figure 5.4.

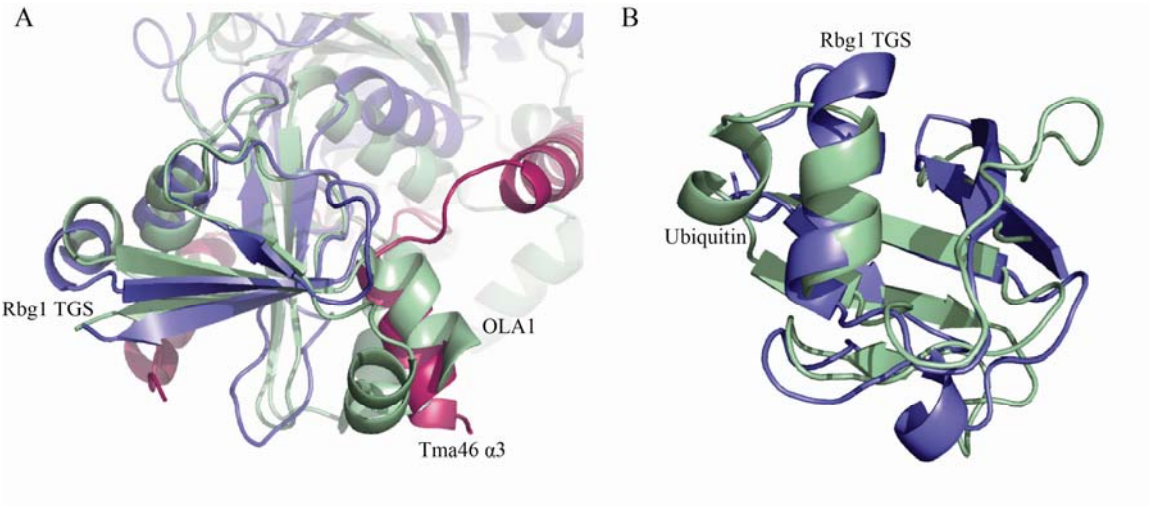


Figure 5.4: Structural alignment of the Rbg1 TGS domain to YchF hOLA1 and ubiquitin.

A. The TGS domain of the $Rbg1_{fl}$-$Tma46_{205-345}$ structure superposed to that from human OLA1 belonging to the YchF subfamily is represented. Rbg1 is coloured in orange, Tma46 in dark red and OLA1 structure is coloured in cyan.

B. The alignment (SSM superpose RMSD of 2.08 Å over 58 residues) of the Rbg1 TGS domain coloured in blue with ubiquitin represented as green.

It is noteworthy to mention here that *A. thaliana* and *P. sativum* Drg2 have a C-terminal extension of around 32 amino acids after the TGS domain in comparison to mammalian Drg2 proteins. This extension is also present in *C. elegans* Drg2, but unlike AtDrg2 and PsDrg2, it is much longer, with about 90 amino acids. The question of whether or not these extensions would form part of the TGS domain or fold into a new domain is interesting.

Two previously unidentified domains were also discovered in Rbg1 by virtue of the structural solution; the HTH and S5D2L domains. Structural similarity searches for the HTH domain gave the Chaperone SYCD (2VGX-B), Female germline-specific tumor suppressor GLD-1 (3K6T-D), NADB-Rossmann superfamily protein among others. In many of these cases, Helix-Turn-Helices are stabilized by a hydrophobic 'zipper' which is also the case in Rbg1 which had one helix (α2) longer than the other as generally seen for Helix-Turn-Helices. HTHs are also commonly involved in dimerization, like the GLD-1

Qua1 HTH which was involved in homodimerization. Rbg1 in complex with the C-terminal fragment of Tma46 eluted as a monomer in size exclusion chromatography and considering that it belongs to the TRAFAC class of proteins which are characterised to function as monomers, it would probably be unlikely that the HTH domain would be involved in dimerising Rbg1. However, crystal contacts of the complex structure do show that the Rbg1 HTH interacts with itself in the symmetry related molecule (HTH′) also including the S5D2L helix suggesting a propensity for interacting with another HTH domain. The interface between the HTH-S5D2L (in molecule A) and HTH-S5D2L′ of the symmetry related molecule (molecule B) comprises roughly of 449.6 $Å^2$ buried surface area as calculated by Pisa.

Mutation analysis reported previously has suggested that Drg1 was binding to the basic helix-loop-helix (bHLH) domain of the oncogenic Tal1 protein and that it might be mediated through interaction with as yet unknown amphipathic helices in Drg1 (Mahajan *et al.,* 1996). Drg1 was also shown to interact *in vitro* with Tal2 and Lyl1. These interacting proteins belong to the family of basic helix-loop-helix (bHLH) transcription factors, whose members also include C-Myc and N-Myc. The bHLH motif is characterized by two alpha helices connected by a loop. In general, transcription factors containing this motif are also dimeric, each with one helix containing basic amino acid residues that facilitate DNA binding. Generally, one helix is smaller than the other and, due to the flexibility of the loop, allows dimerization by folding and packing against another helix. The larger helix typically contained the DNA-binding regions. In another study on *C. albicans* Drg1, it was shown by Yeast-Two-Hybrid to bind to Efg1 which is also a bHLH transcription factor. From these studies, it was speculated that Drg1 in its nucleotide bound/unbound state might be able to sequester (binding and regulating transcription factors thus influencing growth control) transcription factors like Tal11 in the cytoplasm (Chen and Kumamoto 2006).

In yet another study, Drg1 N-terminal residues 1-65 (now known to comprise the HTH domain) were shown to be the only region necessary for binding to MPSK1/STK16 protein (Eswaran *et al.,* 2008). Considering the fact that it was previously unknown of the presence of a domain containing amphipathic helices in Drg1, structural analysis has revealed that MPSK1 reported binding to Drg1 takes place through the HTH domain and the interaction of Drg1 with the above said bHLH factors might be taking place through this newly found N-terminal HTH domain of Drg1.

The novel S5D2L domain of Rbg1 lies adjacent to the HTH domain and is inserted between the conserved GTPase signature elements of the G-domain. It was found to adopt a fold similar to the prokaryotic and eukaryotic ribosomal S5 subunit protein C-terminal domain. This small ribosomal subunit protein paves the entrance of the mRNA channel in the ribosome and is, interestingly, the target of mutations affecting the translational fidelity in the protein synthesis process. These ribosomal ambiguity mutation (RAM) residues of

S5 are also curiously conserved in the Rbg1 S5D2L domain. Ribosomal S5 protein is not a primary rRNA binding protein but does have a surface patch of conserved aromatic and basic residues diagnostic of specific interactions with RNA. The exposed surface of the S5 C-terminal domain formed by the beta sheets has mostly hydrophobic and basic residues, whereas the S5D2L domain of Rbg1, in addition to having basic residues especially lysines, also has non-basic residues like aspartic acid, serine and threonine.

In the ribosome, the subunit protein with which there is evidence of direct interaction for the S5 is the S8 which was shown to form complex in solution (also shown by cross linking studies). However in the structure of the prokaryotic ribosome, the nearby proteins to S5 also include the S4. Visual inspection shows that the C-terminal domain of S5 could have contacts with the S4 protein. The S4 protein as in S5 is involved in translational accuracy and proof reading (Andersson *et al.,* 1986). It is interesting that Rbg1 has two domains, S5D2L and TGS which are structurally similar to ribosomal S5 and S4 proteins respectively (Figure 5.5), raising the question whether DRG proteins could be involved in ensuring accuracy in protein synthesis. Preliminary translational fidelity reporter assays performed in this study unfortunately however did not reveal any such relation.

Rbg1 S5D2L domain is also quite similar to the Elongation Factor-G domain IV. EF-G domain IV, along with domains III and V, is said to mimic the tRNA in the EF-Tu ternary complex in a structural comparison between EF-G and EF-Tu-aminoacyl-tRNA ternary complex (Martemyanov and Gudkov 2000). The fact that DRGs share the presence of a domain which is also found in other translational machinery components like ribosomal protein S5, and domain IV of elongation factors EF-G and eEF-2 suggests further that it might play an important role in the process of translation. The identification of the S5D2L domain in the DRG proteins brings it more phylogenetically closer to the classic translation factor family than any other OBG family member proteins.

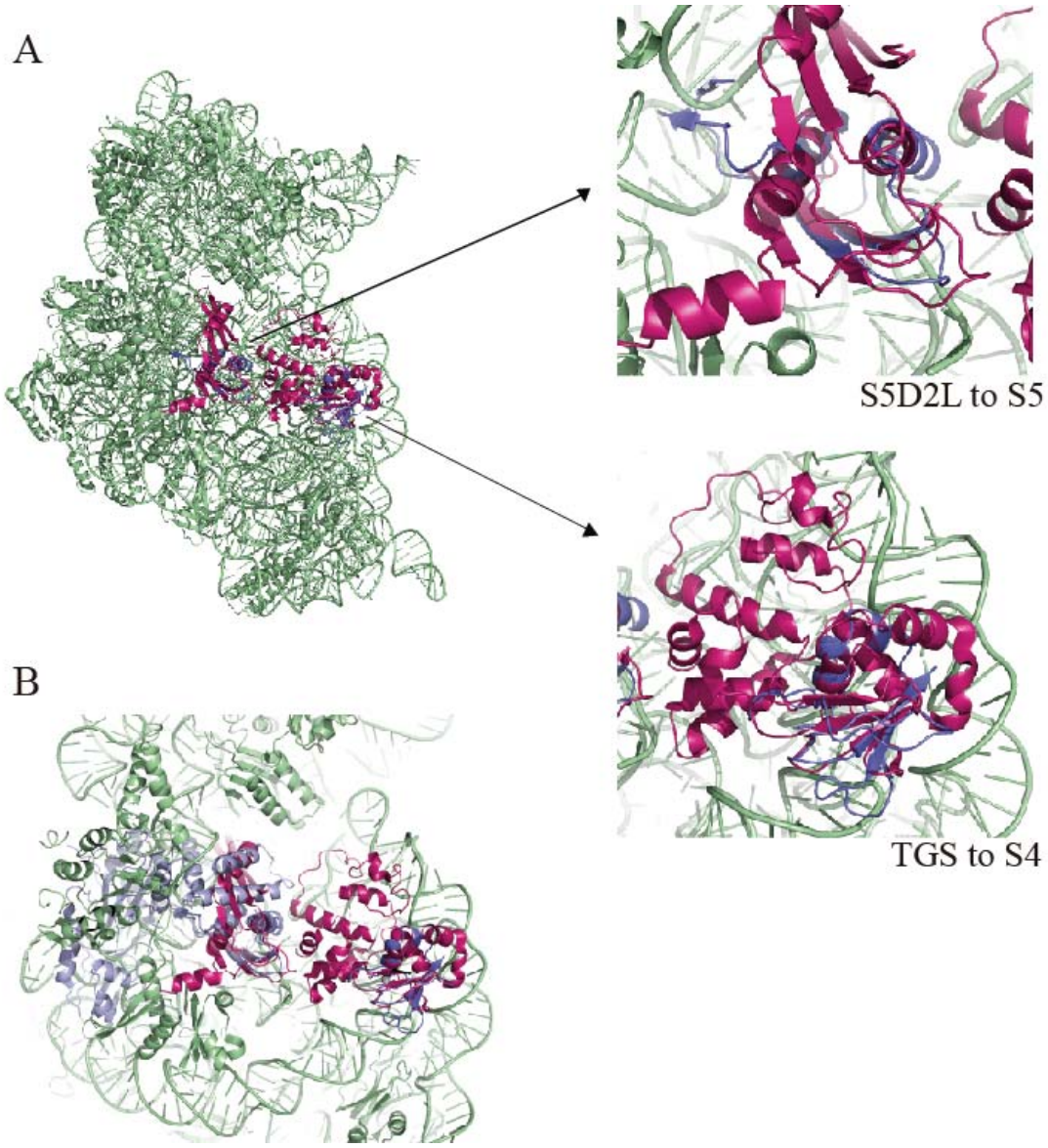


Figure 5.5: Superimpositions of $Rbg1_{fl}$-$Tma46_{205-345}$ to the 30S ribosome.

A. The Thermus thermophilus HB8 30S ribosomal structure (PDB id: 2VQF) is represented in pale green with the ribosomal proteins S4 and S5 which lie adjacent to each other coloured in pink. The S5D2L and TGS domains of Rbg1 are shown in light blue superimposed (SSM) to the S5 and S4 respectively. A zoom in on the domains aligned with the ribosomal proteins is also shown.

B. The $Rbg1_{fl}$-$Tma46_{205-345}$ structure is shown in light blue on the left of the alignment of the S5D2L domain to the S5 ribosomal protein. The figure shows that the orientation of the S5D2L domain in the Rbg1 protein would not allow the simultaneous positioning of the TGS domain over the S4 protein.

As opposed to the other OBG family member proteins, increasing number of evidences are demonstrating that DRGs require the DFRP proteins for its presence and function in the cell and maybe also vice versa. DFRP proteins have been shown in many studies to protect DRGs from polyubiquitination-mediated proteasome degradation, also

affecting its stability and proper expression (Ishikawa *et al.,* 2009, Ishikawa *et al.,* 2005). In our case, $Rbg2_{fl}$ could not be expressed on its own in bacterial expression system and $Rbg1_{fl}$ is not stable. $Tma46_{fl}$ was also found degraded in comparison to the C-terminal fragment as probably the latter is in contact with Rbg1 which prevented it from degradation. Several iterative cycles of construct optimization led to the C-terminal Tma46 fragment in complex with Rbg1 which ultimately crystallised. On the other hand, $Gir2_{fl}$ could be expressed and purified without degradation, probably due to the compact RWD domain in the N-terminal unlike Tma46. Even though the expression of the individual proteins was not straightforward, as complexes the proteins expressed and purified very well in overall.

The fact that the dfrp domain of Tma46 envelops around Rbg1 protein in the structure suggests clearly a protective mechanism from degradation. Also, Tma46 interaction largely occludes parts of the G-domain and TGS domain surface and it was earlier speculated that polyubiquitination could be occurring through the TGS domain by its similarity to ubiquitin. Consistent with the structure, truncated forms of Rbg1 containing only the TGS domain ($Rbg1_{272\text{-}369}$-$Tma46_{154\text{-}345}$) or lacking the TGS domain (($Rbg1_{1\text{-}294}$-$Tma46_{205\text{-}345}$)) still forms complex with Tma46 but shows weaker affinity in an imidazole gradient on a nickel chelating column (data not shown), suggesting that the overall interaction of the C-terminal dfrp fragment of Tma46 with both G-domain and TGS domain of Rbg1 is necessary for a complete stability of the binding. S5D2L or HTH deletion does not affect the binding *in vitro* consistent with the lack of any proximity of Tma46 to these domains in the structure obtained. In a study involving *C. albicans* Drg1, the authors were unable to complement drg1 null mutation with a C-terminal GFP-tagged Drg1 whereas N-terminal myc-tagged Drg1 was biologically active (Chen and Kumamoto 2006). The structural study reveals now that the C-terminal tagging could have probably prevented Dfrp1 from binding to Drg1 while N-terminal tags are on the HLH domain which does not interact with Dfrp1.

The $Tma46_{205\text{-}345}$ dfrp fragment which was crystallized in complex with $Rbg1_{fl}$ contained helices (α1-α4) and a beta strand; α1, α2 and β1 interacting with the G-domain and α3 and α4 interacting with the TGS domain. Helix α2 is quite long and extends from the G-domain with its C-terminal end making contacts with the TGS domain. Rbg1 helix α8 which is in close proximity to G4 and G5 motifs of the G-domain, inserts itself in between Tma46 helices α1 and α2, consisting of a striking pi stacking interaction between Rbg1 α8 and Tma46 α2. We had designed mutations in Tma46 targeting the residues involved in this pi-stacking region and also some residues present in the dfrp domain of Tma46 which are also conserved in Gir2 (yeast Dfrp2 homologue). Mutants $Rbg1_{fl}$ -$Tma46_{205\text{-}345\ G272A\ R273E}$ and $Rbg1_{fl}$ -$Tma46_{205\text{-}345\ W249A\ K250E}$, $Rbg1_{fl}$ -$Tma46_{205\text{-}345\ I241A\ F246A}$ and $Rbg1_{fl}$ -$Tma46_{205\text{-}345\ I241W}$ did not disrupt the binding of Tma46 to Rbg1 clearly due to the extensive surface of interaction between the two proteins.

The secondary structure organization of the $Rbg1_{fl}$ and $Tma46_{205-345}$ proteins in the structure coincided very well to that predicted by Psipred server which predicts secondary structure based on the amino acid sequence. Psipred predicted the helices of the Tma46 dfrp fragment which suggests that it is likely that Tma46 could be forming secondary structures before complex formation with Rbg1.

5.3 DRG-DFRP factors have GTP binding and hydrolytic properties modulated by its component domains

Since early studies, DRG proteins had been shown to bind GTP and GDP (Sazuka *et al.,* 1992, Sommer *et al.,* 1994). However, until now only one report of the characterization of GTP hydrolysis of DRG proteins have been reported (O´Connell, 2009). In these studies, recombinant, refolded, active atDrg1 and atDrg2 from *Arabidopsis thaliana* were shown to bind *in vitro* to both GTP and GDP without the assistance of GAPs or GEFs, unlike Ras-like proteins. AtDrg1 and atDrg2a were shown to have a hydrolysis rate of 0.0009 min^{-1} and 0.0013 min^{-1} respectively. Obg, CgtA, YchF family proteins have also been shown to hydrolyse GTP and rapidly exchange GTP/GDP without the help of GAPs or GEFs (Buglino *et al.,* 2002, Trach and Hoch 1989, Welsh *et al.,* 1994). *Bacillus subtilis* Obg was shown to have a hydrolysis rate of 0.0061 min^{-1} with a K_m of 5.4 μM (Welsh K.M., 1994) and *Caulobacter crescentus* CgtA protein was reported to hydrolyse GTP at a rate of 0.030 min^{-1} (Lin *et al.,* 1999). YchF hOLA1 hydrolysed ATP instead of GTP with a K_{cat} of 0.05 min^{-1} (Koller-Eichhorn *et al.,* 2007).

DRG proteins are similar to translation factors like EF-Tu and EF-G in that they are muti-modular and are capable of hydrolysing GTP. In the present study, we have tried to characterise the GTP binding and hydrolytic properties of the DRGs analysing the importance of each module of the complex on its activity *in vitro*.

$Rbg1_{fl}$-$Tma46_{205-345}$, $Rbg2_{fl}$-$Gir2_{174-238}$ and $Drg1_{fl}$-$Lerepo4_{220-426}$ complexes showed GTP hydrolytic activities (Refer table 4.6) comparable to that of the above said *A. thaliana* DRGs and Obg-YchF proteins. Mutations which are widely used for inactivating GTPase activity are the ones in the G1 motif. Daugeron *et. al.* have previously shown that GTPase activity is necessary for Rbg1-Tma46 complex function *in vivo* and that mutations in some catalytic residues in G1 abolish growth in yeast (Daugeron *et al.,* 2011). Accordingly, the inactive mutant (GFPSVGK**N**) from this study did not show activity in our *in vitro* assay as a complex with Tma46 ($Rbg1_{fl\ GFPSVGKN}$- $Tma46_{205-345}$), serving as a negative control. The $Rbg1_{fl\ GFPSVAMN}$- $Tma46_{205-345}$ mutant abolished any nucleotide binding as measured by the thermal shift assay.

Human Drg1 protein showed reduced GTP binding and hydrolytic properties on its own unlike the complex protein with $Lerepo4_{220-426}$. In the case of the yeast Rbg1 protein,

much decreased hydrolytic properties were observed which could probably be due to the fact that in the absence of DFRP1, Rbg1 was a more unstable protein in our hands than Drg1. O´Connells report have previously shown hydrolytic activity for atDRGs alone and not in complex with DFRP proteins (O'Connell *et al.,* 2009). Interestingly though, in our assay, addition of Lerepo4$_{220\text{-}426}$ to Drg1 just prior to the assay greatly enhanced the catalytic rate of Drg1 (Figure 4.14.C). We speculate that this could either be due to a general stabilization of the DRG1 proteins by DFRP1 proteins or the proximity, as observed from the yeast Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ structure, of the DFRP1 helices to DRG1 helix present near to the G-domain (α8 in Rbg1) directly affecting its GTP activity causing a direct signalling effect of DFRP1 on the DRG1 catalytic site.

Mutations in pi stacking interaction residues or deletion of the α1 helix of Tma46 did not abolish the interaction with Rbg1 and they showed catalytic activity albeit reduced. Although it could not be said that Tma46 is absolutely essential for Rbg1 hydrolytic activity, it might be having a modulating effect. It is curious also that the α2 helix holding the pi stacking residues of Tma46 connects Rbg1 G-domain with the flexible TGS domain raising the question whether this helix is instrumental in transducing signals induced by GTP hydrolysis to the TGS domain for performing a particular function.

TGS and S5D2L domains were found to be not entirely essential for GTPase activity as the protein still hydrolyses GTP but to a lesser degree. The diminution in the hydrolytic rate could again be due to lesser stability of the Rbg1-Tma46 deletion mutant proteins.

5.4 DRG-DFRP complexes are capable of binding RNA non-specifically

Structural analysis shows that the Rbg1 TGS and S5D2L domains lie on either side of the G-domain but forming a curved continuous groove with the distal part of the GTP binding site of the G-domain in the middle. Electrostatic surface potential analysis shows this surface to be predominantly positive (Figure 4.11) which would be electrostatically complementary to RNA, with the curvature providing a perfect complementary surface fit for binding to the ribosome. This electrostatic potential molecular surface is comparatively the most prominently positive surface on the Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ complex. Given the association of DRG-DFRPs to translating ribosomes and a probable RNA binding surface on Rbg1, impelled us to dissect the contribution of the different units of DRG-DFRP on non-specific RNA binding.

In the *in vitro* polyuridylic acid binding studies, Rbg1$_{fl}$-Tma46$_{205\text{-}345}$ complex bound to the homopolymer with sufficient strength and affinity (Section 4.9). The dfrp domain of DFRP1 protein does not seem to bind RNA in that human Dfrp1, as seen with the C-terminal Lerepo4$_{220\text{-}426}$, does not show any binding on its own. Nevertheless, the Zinc

fingers of Dfrp1 does indeed seem to bind RNA as the N-terminal Lerepo4$_{1-220}$ bound to poly(U) in our experiment.

Rbg1/Drg1 on its own also bound non-specifically to RNA. The electrostatic and steric complementarities of the groove suggest that probable binding surface involves the three domains TGS, G-domain and the S5D2L. Although helices α3 and α4 of Tma46 are enveloping TGS domain and residues 283-301 (Tma46 molecule C) were not modelled, we think that the contribution of Tma46 to this potential RNA binding site is minimal based on structural evidence.

Although the TGS domain is speculated to be an RNA binding domain, no direct evidence exists so far that Drg1 could bind to RNA through the TGS domain. Drg1 was shown to bind to polyuridylic acid RNA homopolymer even in the absence of the TGS domain and so it was speculated in this earlier study that the TGS domain might be playing a role in regulating RNA binding or facilitating the access to RNA (Ishikawa *et al.,* 2003). In our hands, yeast Rbg1$_{1-294}$-Tma46$_{205-345}$ lacking the TGS domain also bound RNA. On the other hand, TGS domain does also seem to bind RNA on its own either in the absence (as in the case of Drg1$_{289-367}$) or presence of Tma46 (as in the case of Rbg1$_{272-369}$-Tma46$_{154-345}$). This suggests that the TGS domain can indeed function as an RNA binding domain albeit in combination with additional domains.

S5D2L has the characteristics of a RNA binding domain as most of the members of the Ribosomal protein S5 domain 2-like superfamily are RNA/DNA binding proteins. The S5D2L domain is a 2-layer, α-β domain in which the helices are on one side and the beta sheets on the other which is the case with ribosomal proteins and other RNA-binding domains. However, its topology is unlike any other characterized RNA-binding domains and contains an unusual βαβ left-handed crossover unit. It would be expected that the S5D2L domain would bind RNA, but contrarily, the domain expressed alone did not show any binding to poly(U) in our study. Furthermore, deletion mutants of the S5D2L domain as in Rbg1$_{\Delta 175-243+G}$-Tma46$_{205-345}$ did not abolish binding of the complex to RNA. Although S5D2L did not show any nonspecific RNA binding in our assay, it could probably be mimicking an RNA binding domain or involved in providing specificity to RNA binding.

5.5 Functionality of Drg1 is dependent on its domains and interaction with Dfrp1

The recent study by the laboratory of Dr. Bertrand Séraphin had opened up the possibility of a phenotype assay showing negative yeast growth in a triple deletion mutant of Δ*rbg1*Δ*rbg2*Δ*slh1* (detailed in introduction, (Daugeron *et al.,* 2011). Removal of the TGS domain was shown to be destabilising the Rbg1 protein in this study. Even when overexpressed to restore wild type protein level, there was no complementation of yeast growth by the TGS deletion mutant. However, the TGS deletion used was based only on

the sequence information and probably extended into the G-domain resulting in instability of the mutant protein.

In collaboration with Dr. Séraphin, we sought to analyse the role of each of the domains of Rbg1 on yeast growth based on the structural information now available (results provided in Francis et al, 2012, appendix 8.3). Yeast vectors carrying precise domain deletion mutants of functional HA tagged Rbg1 when transformed in a Δ*rbg1*Δ*rbg2*Δ*slh1* triple mutant strain were analysed for their ability to complement its slow and temperature sensitive growth. The structure of $Rbg1_{fl}$-$Tma46_{205-345}$ complex reveals that S5D2L domain is inserted within the G-domain and hence for the S5D2L domain deletion mutants, one or more glycines were added at the deletion point connecting the two parts of the G-domain to allow sufficient length and flexibility for protein folding ($Rbg1_{\Delta 175-243+G}$ and $Rbg1_{\Delta 171-239+2G}$). The HTH mutant consisted of the deletion of the N-terminal 48 residues of Rbg1 and the TGS mutant consisted of the construct, $Rbg1_{1-293}$.

Although deletion of the entire HTH or S5D2L domain individually did not affect Rbg1 mutant protein expression levels in yeast, the effect on growth were different for the two domain deletions. Removal of HTH domain as was the case also with the deletion of the TGS domain caused the protein to be non-functional at both 30ºC and 37ºC growing conditions whereas with S5D2L deletion, interestingly, the protein was functional at 30ºC but not in a stress condition like 37ºC even though expression levels were the same at both temperatures. This result demonstrated that the S5D2L deletion did not disrupt the protein but the presence of the domain is essential for the full activity of Rbg1 *in vivo*. The two other modules, HTH and TGS domains were found to be indispensable for Rbg1 function.

On the other hand, removing defined Rbg1 interacting helical elements of Tma46 either from the C-terminal end or the N-terminal end suggested that the β strand and α4 helix could be deleted without abrogating function, i.e., the growth of the triple mutant strain *Δgir2Δtma46Δslh1* could be restored to wild type levels on complementation with the mutant vectors. However, deletions of α1, α2 and α3 from either end progressively destabilized the protein resulting concomitantly in a poor complementation with deletions including α2, the most deleterious. Deletion of these helices in Tma46 also resulted in strongly reduced or absent interactions with Rbg1 in co-precipitation assays correlating well with phenotypic interaction, indicating that interaction is synonym of function. The substitution of Tma46 helix alpha 2 with 26 alanines disrupted the interaction with Rbg1 in co-precipitation assays. However, residual interaction was detected with the overexpressed mutant protein and also, on complementation only a slightly reduced growth phenotype was observed at 37ºC. The point residue mutants $Tma46_{205-345\ W249A\ K250E}$ and $Tma46_{205-345\ I241A\ F246A}$ also showed slightly reduced growth phenotype at elevated temperature.

Polysome analyses by sucrose gradient sedimentation of the Rbg1 HTH, TGS and S5D2L domain deletion mutants by Séraphin´s group demonstrated that unlike HTH and S5D2L mutants, deletion of TGS domain disrupted the association of Rbg1 to translating ribosomes. This result would go consistently with the non-specific binding shown by TGS to RNA in our assays. Structural and experimental analyses have suggested that TGS domain is essential for interaction with Tma46 along with the G-domain. It is noteworthy also that Tma46 helix α2 bridges the GTPase activity center with the loosely connected TGS domain. In the context of the DRG factors, these results suggest that the TGS domain whose function was previously unknown, is playing an important role along with Tma46 in recruiting these factors to polysomes also being essential for the functionality of the DRG-DFRP complex.

In order to understand the function of the complex in the protein synthesis mechanism, initial assays were carried out in which slight variations in the proper stop codon recognition in the termination step of translation was detected between the wild type *S. cerevisieae* strain and that containing deletions of the three genes *rbg1*, *rbg2* and *slh1* from their genome. This result supports the role of the DRG factors in translation. Further experimental evidence could help shed more light into the exact mechanism by which the DRGs are associated to protein synthesis.

5.6 Structure-based dissection of Rbg1-Tma46/Rbg2-Gir2 and Rbg1-Gir2 interactions suggest differences

Early studies have demonstrated that Drg1 and Drg2 protein expression levels are greatly reduced in the absence of their corresponding interaction partners, Dfrp1 and Dfrp2. Co-transfection of Dfrp2, however has been shown to increase both Drg1 and Drg2 expression levels whereas Dfrp1 increased only Drg1 and not Drg2 expression (Ishikawa *et al.,* 2005). Wout *et al.* have also reported that yeast Drg1 interacts with Dfrp2, which in turn binds to Gcn1, the resulting complex probably able to associate to polysomes (Wout *et al.,* 2009). Consistently, in our RNA binding studies, Gir2 could not interact with polyuridylic acid on its own unlike Dfrp1.

Drg1 and Drg2 are evolutionarily very close to each other. Dfrp1 and Dfrp2, on the other hand, share little similarity in their N-terminal, being only vaguely homologous in the C-terminal dfrp fragment that is involved in interactions with the DRG factors. Some of the residue differences between Rbg1 and Rbg2 in the interface of Rbg1-Tma46 are contributed by hydrogen bonds with the main chain atoms, thus the change in side chains would not affect the interaction. Yet other interactions between Rbg1 and Tma46 involved conserved amino acid residues.

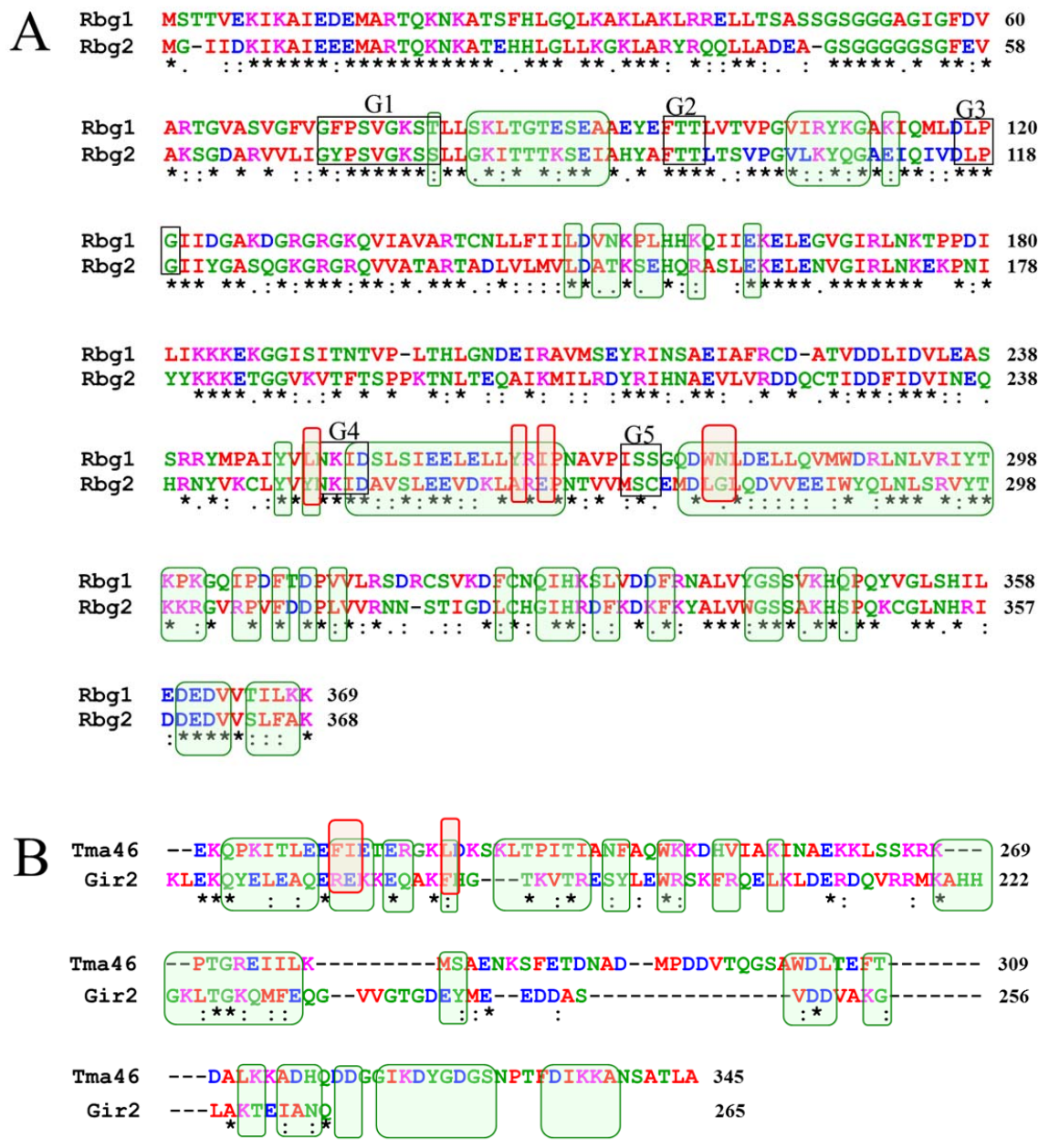


Figure 5.6: Sequence alignment showing differences in interaction residues in Rbg1/Rbg2 and Tma46/Gir2.

A. The residues of Rbg1 which interact with Tma46 are given inside the green boxes. The five G-motifs are indicated and the prominent amino acid residue differences between Rbg1 and Rbg2 are given as a red box.

B. The alignment of the C-terminal dfrp part of Tma46 and Gir2 is shown with the dissimilar residues shown in red boxes.

The amino acid differences that do exist are in the interaction area with Rbg1 at the G-domain. This includes hydrophobic interactions contributed by Rbg1: Leu249, Ile266

and Tma46: Phe225, Leu222 which get replaced in Rbg2 by Tyr249, Glu266 and Gir2 Arg178, Ala175 respectively. Arg178 of Gir2 could form hydrogen bonds with Rbg2 Tyr249 or salt bridge with Rbg2 Glu266. In overall, while the interaction in this region involving the α1 and α2 helices of Tma46 seems predominantly hydrophobic for Rbg1, in the case of Rbg2-Gir2 could be mostly hydrophilic.

Interestingly, the most prominent difference existing in the interface of the two binding proteins is the pi stacking region between Rbg1 α8/α9 and Tma46 α2, where five ring residues stack against each other. While these aromatic residues contributing to the stacking interaction is conserved in Rbg1, Tma46 and Gir2, they are not present in Rbg2 which suggests that the possible interaction of Rbg1 with Gir2 might be fortified by this type of interaction. It is interesting to think that this could also be the reason why Rbg2 is not interacting with Tma46 as evidenced in previous experiments.

Although Dfrp2 interact with Drg1 *in vitro*, studies suggest that in higher organisms, Drg1 and Drg2 complexes do not exist together. Different expression patterns of *drg1* and *drg2* mRNA transcripts were seen by Northern blotting and whole mount in situ hybridization in *X. laevis* adult and embryonic tissues (Ishikawa *et al.,* 2003). Dfrp1 expression was similar to Drg1 and not to Drg2 suggesting regulation between the two. Spatial expression patterns and protein levels of Drg1 and Drg2 were also different in plants (Stafstrom 2008). Also, in yeast, TAP affinity purifications have detected the two complexes, Rbg1-Tma46 and Rbg2-Gir2 independently with no visible cross-talk (Daugeron *et al.,* 2011). We speculate that the functionality of the Rbg1-Gir2 complex seen in our experiment could be attributed to the occurrence more commonly of functionally redundant factors in the unicellular yeast. In higher organisms, however, Drg1-Dfrp1 and Drg2-Dfrp2 might have independent functions and/or Drg1 could have independent roles either bound to Dfrp1 or Dfrp2.

Conclusions

6. Conclusions

1. The Developmentally Regulated GTPase family of proteins belonging to the TRAFAC class are monomeric and multi-modular, containing three domains in addition to the core GTP binding center.
2. We have identified a novel domain in DRG factors which is structurally similar to Helix-Turn-Helices known to have propensities towards protein/DNA binding.
3. We have found a previously unidentified domain in Drg1 with unknown function structurally related to the Ribosomal S5 Domain2-like superfamily (S5D2L).
4. The TGS and G-domain of Drg1 are involved in interactions with its binding partner, the Developmentally Regulated GTPase family regulatory protein 1 (Dfrp1).
5. Structurally uncharacterised Dfrp1 C-terminal domain is now shown to be non-globular and interacts with Drg1 forming an extended interface. This mode of interaction provides a rationale to the previously reported DFRP mediated post-transcriptional regulation stabilisation and prevention from degradation of DRGs, also modulating its GTPase activity.
6. Isolated DRG factors have non-specific RNA binding properties mediated by the TGS and GTP binding domains.
7. Rbg1 TGS domain and Tma46 are essential for recruitment of Rbg1 to translating ribosomes. Our structural and experimental evidence demonstrates that this takes place through a direct interaction between Tma46 and Rbg1 TGS domain.
8. Since all OBG family components present GTPase activity, the specific biological functions of a given member are determined by the non G-domains. To this respect, the DRG family is phylogenetically closer to the translation factor family than any other OBG family member thus providing structural support for its role in translation.

Conclusiones

1. La familia de proteínas DRG perteneciente a la clase TRAFAC son monoméricas y multimodulares que contienen 3 dominios además del de unión a GTP.
2. Hemos identificado un nuevo dominio en los factores DRG que es estructuralmente similar a los Helice-giro-hélice conocidos por sus propiedades de unión proteína/ADN.
3. Hemos encontrado en DRG1 un dominio previamente desconocido de función desconocida y estructuralmente relacionado con la superfamilia "Ribosomal S5 Domain2-like" (S5D2L).
4. Los dominios TGS y G de Drg1 están involucrados en interacciones con su ligando celular, Dfrp1.
5. El dominio C-terminal de Dfrp1 estructuralmente desconocido con anterioridad se muestra ahora como un dominio no globular e interacciona con Drg1 formando una extensa interfaz.
6. Los factores DRG aislados tiene propiedades de unión no específica al ARN mediante los dominios TGS y G.
7. Tanto el dominio TGS de Rbg1 como Tma46 son esenciales en el reclutamiento de Rbg1 a ribosomas en traducción. Nuestras evidencias estructurales y experimentales demuestran que este fenómeno tiene lugar mediante una interacción directa entre Tma46 y el dominio TGS de Rbg1.
8. Puesto que todos los componentes de la familia OBG poseen actividad GTPasa, las funciones biológicas de un miembro dado vienen determinadas por otros dominios distintos al G. A este respecto la familia DRG está filogenéticamente mas próxima a la familia de factores de traducción que cualquier otro miembro de la familia OBG proporcionando apoyo estructural para su papel en traducción.

References

7. References

Alves, V.S. and Castilho, B.A., "Gir2 is an intrinsically unstructured protein that is present in Saccharomyces cerevisiae as a group of heterogeneously electrophoretic migrating forms", *Biochem Biophys Res Commun*, 2005, 332, (2), 450-455.

Alves, V.S., Pimenta, D.C., Sattlegger, E. and Castilho, B.A., "Biophysical characterization of Gir2, a highly acidic protein of Saccharomyces cerevisiae with anomalous electrophoretic behavior", *Biochem Biophys Res Commun*, 2004, 314, (1), 229-234.

Andersson, D.I., Andersson, S.G. and Kurland, C.G., "Functional interactions between mutated forms of ribosomal proteins S4, S5 and S12", *Biochimie*, 1986, 68, (5), 705-713.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A., "Electrostatics of nanosystems: application to microtubules and the ribosome", *Proc Natl Acad Sci U S A*, 2001, 98, (18), 10037-10041.

Baudin-Baillieu, A., Guillemet, E., Cullin, C. and Lacroute, F., "Construction of a yeast strain deleted for the TRP1 promoter and coding region that enhances the efficiency of the polymerase chain reaction-disruption method", *Yeast*, 1997, 13, (4), 353-356.

Baykov, A.A., Evtushenko, O.A. and Avaeva, S.M., "A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay", *Anal Biochem*, 1988, 171, (2), 266-270.

Bourne, H.R., Sanders, D.A. and McCormick, F., "The GTPase superfamily: a conserved switch for diverse cell functions", *Nature*, 1990, 348, (6297), 125-132.

Bourne, H.R., Sanders, D.A. and McCormick, F., "The GTPase superfamily: conserved structure and molecular mechanism", *Nature*, 1991, 349, (6305), 117-127.

Bradford, M.M., "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding", *Anal Biochem*, 1976, 72, 248-254.

Buglino, J., Shen, V., Hakimian, P. and Lima, C.D., "Structural and biochemical analysis of the Obg GTP binding protein", *Structure*, 2002, 10, (11), 1581-1592.

Campos-Olivas, R., Sanchez, R., Torres, D. and Blanco, F.J., "Backbone assignment of the 98 kDa homotrimeric yeast PCNA ring", *J Biomol NMR*, 2007, 38, (2), 167.

Capalbo, G., Muller-Kuller, T., Dietrich, U., Hoelzer, D., Ottmann, O.G. and Scheuring, U.J., "Inhibition of X4-tropic HIV type 1 replication by knockdown of the cellular protein LEREPO4", *AIDS Res Hum Retroviruses*, 2010, 26, (10), 1155-1161.

Catimel, B., Schieber, C., Condron, M., Patsiouras, H., Connolly, L., Catimel, J., Nice, E.C., Burgess, A.W. and Holmes, A.B., "The PI(3,5)P2 and PI(4,5)P2 interactomes", *J Proteome Res*, 2008, 7, (12), 5295-5313.

Catimel, B., Yin, M.X., Schieber, C., Condron, M., Patsiouras, H., Catimel, J., Robinson, D.E., Wong, L.S., Nice, E.C., Holmes, A.B. and Burgess, A.W., "PI(3,4,5)P3 Interactome", *J Proteome Res*, 2009, 8, (7), 3712-3726.

CCP4, C.C.P., "The CCP4 suite: programs for protein crystallography", *Acta Crystallogr D Biol Crystallogr*, 1994, 50, (Pt 5), 760-763.

Chen, X. and Kumamoto, C.A., "A conserved G protein (Drg1p) plays a role in regulation of invasive filamentation in Candida albicans", *Microbiology*, 2006, 152, (Pt 12), 3691-3700.

Daugeron, M.C., Prouteau, M., Lacroute, F. and Seraphin, B., "The highly conserved eukaryotic DRG factors are required for efficient translation in a manner redundant with the putative RNA helicase Slh1", *Nucleic Acids Res*, 2011, 39, (6), 2221-2233.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S. and Richardson, D.C., "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids", *Nucleic Acids Res*, 2007, 35, (Web Server issue), W375-383.

Dean, D. and Nomura, M., "Feedback regulation of ribosomal protein gene expression in Escherichia coli", *Proc Natl Acad Sci U S A*, 1980, 77, (6), 3590-3594.

Doublie, S., "Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems", *Methods Mol Biol*, 2007, 363, 91-108.

Emsley, P. and Cowtan, K., "Coot: model-building tools for molecular graphics", *Acta Crystallogr D Biol Crystallogr*, 2004, 60, (Pt 12 Pt 1), 2126-2132.

Eswaran, J., Bernad, A., Ligos, J.M., Guinea, B., Debreczeni, J.E., Sobott, F., Parker, S.A., Najmanovich, R., Turk, B.E. and Knapp, S., "Structure of the human protein kinase MPSK1 reveals an atypical activation loop architecture", *Structure*, 2008, 16, (1), 115-124.

Fleischer, T.C., Weaver, C.M., McAfee, K.J., Jennings, J.L. and Link, A.J., "Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes", *Genes Dev*, 2006, 20, (10), 1294-1307.

Fortelle, E.d.L. and Bricogne, G., "Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods", *Methods in Enzymology*, 1997, 276, 472-494.

Garcia-Ranea, J.A. and Valencia, A., "Distribution and functional diversification of the ras superfamily in Saccharomyces cerevisiae", *FEBS Lett*, 1998, 434, (3), 219-225.

Gregory, R.C., Lord, K.A., Panek, L.B., Gaines, P., Dillon, S.B. and Wojchowski, D.M., "Subtraction cloning and initial characterization of novel epo-immediate response genes", *Cytokine*, 2000, 12, (7), 845-857.

Heese, K., Nagai, Y. and Sawada, T., "Nerve growth factor (NGF) induces mRNA expression of the new transcription factor protein p48ZnF", *Exp Mol Med*, 2004, 36, (2), 130-134.

Hoenig, M., Lee, R.J. and Ferguson, D.C., "A microtiter plate assay for inorganic phosphate", *J Biochem Biophys Methods*, 1989, 19, (2-3), 249-251.

Holm, L. and Rosenstrom, P., "Dali server: conservation mapping in 3D", *Nucleic Acids Res*, 2010, 38, (Web Server issue), W545-549.

Ishikawa, K., Akiyama, T., Ito, K., Semba, K. and Inoue, J., "Independent stabilizations of polysomal Drg1/Dfrp1 complex and non-polysomal Drg2/Dfrp2 complex in mammalian cells", *Biochem Biophys Res Commun*, 2009, 390, (3), 552-556.

Ishikawa, K., Azuma, S., Ikawa, S., Morishita, Y., Gohda, J., Akiyama, T., Semba, K. and Inoue, J., "Cloning and characterization of Xenopus laevis drg2, a member of the developmentally regulated GTP-binding protein subfamily", *Gene*, 2003, 322, 105-112.

Ishikawa, K., Azuma, S., Ikawa, S., Semba, K. and Inoue, J., "Identification of DRG family regulatory proteins (DFRPs): specific regulation of DRG1 and DRG2", *Genes Cells*, 2005, 10, (2), 139-150.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. and Sakaki, Y., "A comprehensive two-hybrid analysis to explore the yeast protein interactome", *Proc Natl Acad Sci U S A*, 2001, 98, (8), 4569-4574.

Jakobs, A., Himstedt, F., Funk, M., Korn, B., Gaestel, M. and Niedenthal, R., "Ubc9 fusion-directed SUMOylation identifies constitutive and inducible SUMOylation", *Nucleic Acids Res*, 2007, 35, (17), e109.

Kabsch, W., Gast, W.H., Schulz, G.E. and Leberman, R., "Low resolution structure of partially trypsin-degraded polypeptide elongation factor, EF-TU, from Escherichia coli", *J Mol Biol*, 1977, 117, (4), 999-1012.

Koller-Eichhorn, R., Marquardt, T., Gail, R., Wittinghofer, A., Kostrewa, D., Kutay, U. and Kambach, C., "Human OLA1 defines an ATPase subfamily in the Obg family of GTP-binding proteins", *J Biol Chem*, 2007, 282, (27), 19928-19937.

Kramer, E.B. and Farabaugh, P.J., "The frequency of translational misreading errors in E. coli is largely determined by tRNA competition", *RNA*, 2007, 13, (1), 87-96.

Kramer, E.B., Vallabhaneni, H., Mayer, L.M. and Farabaugh, P.J., "A comprehensive analysis of translational missense errors in the yeast Saccharomyces cerevisiae", *RNA*, 2010, 16, (9), 1797-1808.

Krissinel, E. and Henrick, K., "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions", *Acta Crystallogr D Biol Crystallogr*, 2004, 60, (Pt 12 Pt 1), 2256-2268.

Kubota, H., Sakaki, Y. and Ito, T., "GI domain-mediated association of the eukaryotic initiation factor 2alpha kinase GCN2 with its activator GCN1 is required for general amino acid control in budding yeast", *J Biol Chem*, 2000, 275, (27), 20243-20246.

Kumar, S., Iwao, M., Yamagishi, T., Noda, M. and Asashima, M., "Expression of GTP-binding protein gene drg during Xenopus laevis development", *Int J Dev Biol*, 1993, 37, (4), 539-546.

Langer, G., Cohen, S.X., Lamzin, V.S. and Perrakis, A., "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7", *Nat Protoc*, 2008, 3, (7), 1171-1179.

Lanzetta, P.A., Alvarez, L.J., Reinach, P.S. and Candia, O.A., "An improved assay for nanomole amounts of inorganic phosphate", *Anal Biochem*, 1979, 100, (1), 95-97.

Leipe, D.D., Wolf, Y.I., Koonin, E.V. and Aravind, L., "Classification and evolution of P-loop GTPases and related ATPases", *J Mol Biol*, 2002, 317, (1), 41-72.

Li, B. and Trueb, B., "DRG represents a family of two closely related GTP-binding proteins", *Biochim Biophys Acta*, 2000, 1491, (1-3), 196-204.

Lin, B., Covalle, K.L. and Maddock, J.R., "The Caulobacter crescentus CgtA protein displays unusual guanine nucleotide binding and exchange properties", *J Bacteriol*, 1999, 181, (18), 5825-5832.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S. and Richardson, D.C., "Structure validation by Calpha geometry: phi,psi and Cbeta deviation", *Proteins*, 2003, 50, (3), 437-450.

Macara, I.G., Lounsbury, K.M., Richards, S.A., McKiernan, C. and Bar-Sagi, D., "The Ras superfamily of GTPases", *FASEB J*, 1996, 10, (5), 625-630.

Mahajan, M.A., Park, S.T. and Sun, X.H., "Association of a novel GTP binding protein, DRG, with TAL oncogenic proteins", *Oncogene*, 1996, 12, (11), 2343-2350.

Martemyanov, K.A. and Gudkov, A.T., "Domain III of elongation factor G from Thermus thermophilus is essential for induction of GTP hydrolysis on the ribosome", *J Biol Chem*, 2000, 275, (46), 35820-35824.

McCarthy, A.A., Brockhauser, S., Nurizzo, D., Theveneau, P., Mairs, T., Spruce, D., Guijarro, M., Lesourd, M., Ravelli, R.B. and McSweeney, S., "A decade of user operation on the macromolecular crystallography MAD beamline ID14-4 at the ESRF", *J Synchrotron Radiat*, 2009, 16, (Pt 6), 803-812.

Mizushima, S. and Nomura, M., "Assembly mapping of 30S ribosomal proteins from E. coli", *Nature*, 1970, 226, (5252), 1214.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J., "Refinement of macromolecular structures by the maximum-likelihood method", *Acta Crystallogr D Biol Crystallogr*, 1997, 53, (Pt 3), 240-255.

Nelson, B.J., Maas, K.J., L., D.J. and Stafstrom, J.P., "Association of DRG1 and DRG2 with ribosomes from pea, arabidosis and yeast", *Int J Plant Sci*, 2009, 170, (7), 834-844.

O'Connell, A., Robin, G., Kobe, B. and Botella, J.R., "Biochemical characterization of Arabidopsis developmentally regulated G-proteins (DRGs)", *Protein Expr Purif*, 2009, 67, (2), 88-95.

Pape, T. and Schneider, T.R., "HKL2MAP: a graphical user interface for phasing with SHELX programs", *J Appl Cryst*, 2004, 37, 843-844.

Pitossi, F., Blank, A., Schroder, A., Schwarz, A., Hussi, P., Schwemmle, M., Pavlovic, J. and Staeheli, P., "A functional GTP-binding motif is necessary for antiviral activity of Mx proteins", *J Virol*, 1993, 67, (11), 6726-6732.

Rudolph, M.G., Wittinghofer, A. and Vetter, I.R., "Nucleotide binding to the G12V-mutant of Cdc42 investigated by X-ray diffraction and fluorescence spectroscopy: two different nucleotide states in one crystal", *Protein Sci*, 1999, 8, (4), 778-787.

Salas-Marco, J. and Bedwell, D.M., "Discrimination between defects in elongation fidelity and termination efficiency provides mechanistic insights into translational readthrough", *J Mol Biol*, 2005, 348, (4), 801-815.

Savvides, S.N., Yeo, H.J., Beck, M.R., Blaesing, F., Lurz, R., Lanka, E., Buhrdorf, R., Fischer, W., Haas, R. and Waksman, G., "VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion", *EMBO J*, 2003, 22, (9), 1969-1980.

Sazuka, T., Kinoshita, M., Tomooka, Y., Ikawa, Y., Noda, M. and Kumar, S., "Expression of DRG during murine embryonic development", *Biochem Biophys Res Commun*, 1992, 189, (1), 371-377.

Sazuka, T., Tomooka, Y., Ikawa, Y., Noda, M. and Kumar, S., "DRG: a novel developmentally regulated GTP-binding protein", *Biochem Biophys Res Commun*, 1992, 189, (1), 363-370.

Schenker, T., Lach, C., Kessler, B., Calderara, S. and Trueb, B., "A novel GTP-binding protein which is selectively repressed in SV40 transformed fibroblasts", *J Biol Chem*, 1994, 269, (41), 25447-25453.

Scheuring, U.J., Corbeil, J., Mosier, D.E. and Theofilopoulos, A.N., "Early modification of host cell gene expression induced by HIV-1", *AIDS*, 1998, 12, (6), 563-570.

Schrodinger, LLC The PyMOL Molecular Graphics System, Version 1.3r1.(2010).

Sheldrick, G.M., "A short history of SHELX", *Acta Crystallogr A*, 2008, 64, (Pt 1), 112-122.

Sommer, K.A., Petersen, G. and Bautz, E.K., "The gene upstream of DmRP128 codes for a novel GTP-binding protein of Drosophila melanogaster", *Mol Gen Genet*, 1994, 242, (4), 391-398.

Sreenath, H.K., Bingman, C.A., Buchan, B.W., Seder, K.D., Burns, B.T., Geetha, H.V., Jeon, W.B., Vojtik, F.C., Aceti, D.J., Frederick, R.O., Phillips, G.N., Jr. and Fox, B.G., "Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium", *Protein Expr Purif*, 2005, 40, (2), 256-267.

Stafstrom, J.P., "Expression patterns of Arabidopsis DRG genes: promoter-gus fusions, quantitative real-time PCR, and patterns of protein accumulation in response to environmental stresses", *Int J Plant Sci*, 2008, 169, (8), 1046-1056.

Staker, B.L., Korber, P., Bardwell, J.C. and Saper, M.A., "Structure of Hsp15 reveals a novel RNA-binding motif", *EMBO J*, 2000, 19, (4), 749-757.

Studier, F.W., "Protein production by auto-induction in high density shaking cultures", *Protein Expr Purif*, 2005, 41, (1), 207-234.

Sugahara, M., Asada, Y., Ayama, H., Ukawa, H., Taka, H. and Kunishima, N., "Heavy-atom Database System: a tool for the preparation of heavy-atom derivatives of protein crystals based on amino-acid sequence and crystallization conditions", *Acta Crystallogr D Biol Crystallogr*, 2005, 61, (Pt 9), 1302-1305.

Trach, K. and Hoch, J.A., "The Bacillus subtilis spo0B stage 0 sporulation operon encodes an essential GTP-binding protein", *J Bacteriol*, 1989, 171, (3), 1362-1371.

Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S. and Rothberg, J.M., "A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae", *Nature*, 2000, 403, (6770), 623-627.

Verstraeten, N., Fauvart, M., Versees, W. and Michiels, J., "The universally conserved prokaryotic GTPases", *Microbiol Mol Biol Rev*, 2011, 75, (3), 507-542.

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J., "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold", *EMBO J*, 1982, 1, (8), 945-951.

Welsh, K.M., Trach, K.A., Folger, C. and Hoch, J.A., "Biochemical characterization of the essential GTP-binding protein Obg of Bacillus subtilis", *J Bacteriol*, 1994, 176, (23), 7161-7168.

Wittinghofer, A. and Vetter, I.R., "Structure-function relationships of the G domain, a canonical switch motif", *Annu Rev Biochem*, 2011, 80, 943-971.

Wolf, Y.I., Aravind, L., Grishin, N.V. and Koonin, E.V., "Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events", *Genome Res*, 1999, 9, (8), 689-710.

Wout, P.K., Sattlegger, E., Sullivan, S.M. and Maddock, J.R., "Saccharomyces cerevisiae Rbg1 protein and its binding partner Gir2 interact on Polyribosomes with Gcn1", *Eukaryot Cell*, 2009, 8, (7), 1061-1071.

Zhao, X.F. and Aplan, P.D., "SCL binds the human homologue of DRG in vivo", *Biochim Biophys Acta*, 1998, 1448, (1), 109-114.

Appendices

8. Appendix

8.1 Abbreviations

α	Alpha helix
β	Beta strand
6His	6x Histidine tag
APBS	Adaptive Poisson-Boltzmann Solver
AMPPCP	Adenosine-5′-[(β,γ)-methyleno]triphosphate
AU	Asymmetric unit
bHLH	Basic helix-loop-helix
BSA	Bovine serum albumin
CC	Correlation coefficient
CCD	Charge-coupled device
CCP4	Collaborative Computational Project Number 4
C-terminal	Carboxy terminal
DFRP	DRG family regulatory protein
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DRG	Developmentally regulated GTP binding protein
DTT	Dithiothreitol
EF	Elongation factor
FF	Fast flow
fl	Full length
GDP	Guanosine 5′- diphosphate
GHMP	Galacto-, homoserine-, mevalonate-, phosphomevalonate- kinases
GIR2	Genetically interacting with ribosomal genes 2
GTP	Guanosine 5′-triphosphate
GTPGS	Guanosine 5′- O-[gamma-thio]triphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kBT	Product of Boltzmann constant (kB) and absolute temperature (T)
K_{cat}	Catalytic constant
K_m	Michaelis-Menten constant
LB	Luria bertani
LEREPO4	Likely ortholog of mouse immediate early response erythropoietin 4
LUCA	Last universal common ancestor
pI	Isoelectric point
MR	Molecular replacement
NCS	Non-crystallographic symmetry
Ni-NTA	Nickel-nitriloacetic acid
NMR	Nuclear magnetic resonance

N-terminal	Amino terminal
OD_{600nm}	Optical density at 600 nm
ON	Overnight
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PGK	Phosphoglycerate kinase-1 promoter
PolyU	Poly uridylic acid
RBG	Ribosome binding GTPase
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RWD	Ring finger, WD repeat and DEAD-like
RWDD1	RWD domain containing protein 1
S	Svedberg unit (sedimentation)
S5	Small ribosomal subunit protein 5
S5D2L	S5 domain 2-like
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SelMet	Selenomethionine
SIRAS	Single isomorphous replacement with anomalous scattering
SLH1	SKI2-like helicase 1
SRF	Self-rotation function
SHARP	Statistical heavy-atom refinement and phasing
SUMO	Small ubiquitin-like modifier
TAL	T-cell acute lympoblastic leukemia protein
TGS	threonyl-tRNA synthetases (ThrRSs), DRG (GTP-binding proteins) and Guanosine polyphosphate phosphohydrolases/ synthetases (SpoT/RelA)
TLS	Translation-libration-screw
Tm	Melting temperature
TMA46	Translation machinery associated protein 46
TRAFAC	Translation Factors
UV	Ultra violet
V_{max}	Maximal velocity

8.2 Recipes

20 x NPS solution (mol/litre):

- 0.5 M $(NH_4)_2SO_4$
- 1 M KH_2PO_4
- 1 M Na_2HPO_4

50 x 5052 solution:

- 25% g glycerol
- 2.5% glucose
- 10% α-lactose

10,000x trace metals solution:

- 0.05 M $FeCl_3.6H_2O$ dissolved in ~0.1M HCl
- 0.01 M $CaCl_2$
- 0.01 M $MnCl_2.4H_2O$
- 0.01 M $ZnSO_4.7H_2O$
- 0.002 M $CoCl_2.6H_2O$
- 0.002 M $CuCl_2.2H_2O$
- 0.002 M $NiCl_2.6H_2O$
- 0.002 M $Na_2MoO_4.5H_2O$
- 0.002 M $Na_2SeO_3.5H_2O$
- 0.002 M H_3BO_3

1000x vitamins solution:

- 0.2 mM Nicotinic acid
- 0.2 mM Pyridoxine-HCl
- 0.2 mM Thiamine-HCl
- 0.2 mM p-aminobenzoic acid
- 0.2 mM Pantothenate
- 5 µM Folic acid
- 5 µM Riboflavin
- 0.2 mM Vitamin B_{12}

50x amino acid mixture:

- 1% Sodium glutamate
- 1% Lysine-HCl
- 1% Arginine-HCl
- 1% Histidine-HCl
- 1% Free acid of aspartic acid
- 1% Zwitterionic forms of alanine
- 1% Praline
- 1% Glycine
- 1% Threonine
- 1% Serine
- 1% Glutamine
- 1% Asparagine
- 1% Valine
- 1% Leucine
- 1% Isoleucine
- 1% Phenylalanine
- 1% Tryptophan

ZY

- 1% N-Z-amine-AS (or tryptone)
- 0.5% Yeast extract

Rbg1-Tma46 dimer structure reveals new functional domains and their role in polysome recruitment

Journal:	*Nucleic Acids Research*
Manuscript ID:	Draft
Manuscript Type:	1 Standard Manuscript
Key Words:	GTPase, translation, TGS domain, Ribosomal protein S5 domain 2-like superfamily, polysome association

Rbg1-Tma46 dimer structure reveals new functional domains and their role in polysome recruitment

Sandrea M. Francis[1*], María-Eugenia Gas[2], Marie-Claire Daugeron[3,4], Jeronimo Bravo[1], Bertrand Séraphin[2,3]

[1] Instituto de Biomedicina de Valencia (IBV-CSIC), Calle Jaime Roig, 11, Valencia E-46010, Spain

[2] Equipe Labellisée La Ligue, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM) U964/Centre National de Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 67404 Illkirch, France.

3 Equipe Labellisée La Ligue, CGM, CNRS UPR3404, 1 Avenue de la Terrasse, 91198 Gif sur Yvette Cedex, France

[4]Université Paris-Sud 11, Orsay, F-91405

Correspondence to: Jeronimo Bravo, Instituto de Biomedicina de Valencia (IBV), Calle Jaime Roig, 11, Valencia – 46010, España. Tel: +34 96 339 1760; Fax: +34 96 369 0800; E-mail: jbravo@ibv.csic.es, and Bertrand Séraphin, IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Illkirch F-67400, France Tel +33 (0)3 88 65 33 36; Fax +33 (0)3 88 65 33 37; E-mail: seraphin@igbmc.fr.

Keywords: GTPase, translation, TGS domain, Ribosomal protein S5 domain 2-like superfamily, polysome association.

Running title: Structure and function of the Rbg1-Tma46 dimer

Abstract

DRG proteins are highly conserved GTPases that associate with DFRP proteins. The resulting complexes have recently been shown to participate in eukaryotic translation. The structure of the Rbg1 GTPase, a yeast DRG protein, in complex with the C-terminal region of its DFRP partner, Tma46, was solved by X-ray diffraction. These data reveal that DRG proteins are multimodular factors with three additional globular domains, HTH, S5D2L and TGS, packing against the GTPase platform. Surprisingly, one of these new domains is inserted in the middle of the GTPase sequence. In contrast, the region of Tma46 interacting with Rbg1 adopts an extended conformation typical of intrinsically unstructured proteins that contacts the GTPase and TGS domains. Functional analyses demonstrate that the various domains of Rbg1, as well as Tma46, modulate the GTPase activity of Rbg1 and contribute to the function of this complex *in vivo*. Dissecting the role of the different domains revealed that the Rbg1 TGS domain is essential for the recruitment of this factor in polysomes, supporting further the implication of these conserved factors in translation.

Introduction

GTPases form a large family of universally represented proteins that have been involved in many cellular functions. Phylogenetic analyses have demonstrated that GTPases organize themselves in two distinct classes (1). The best-known branch was named TRAFAC as it contains GTPases involved in translation (TRAnslation FACtors). Besides translation factors, this branch also encompasses the well-known trimeric GTPases involved in signal transduction, septins and the RAS subfamily of GTPases. These proteins are characterized by the presence of a GTP binding domain (G-domain) that contains five characteristic motifs, G1 (Walker A/P-loop, GxxxxGKS) responsible for binding of α- and β-phosphate groups of the nucleotide, G2 (Switch I) that binds Mg^{2+}, G3 (Walker B/Switch II, DxxG) that binds to Mg^{2+} and interacts with the nucleotide γ-phosphate, G4 (N/TKxD) where K and D bind directly to the nucleotide, and G5 (SAH) involved in guanine base recognition. Small G-proteins have been extensively characterized and found to act as important molecular switches through changes in conformation related to the presence and nature of the bound nucleotide (none, GDP, GTP). In particular, the conformational changes occurring as a result of GTP hydrolysis has been shown to transduce cellular signals to downstream effectors mainly through changes in switch I (G2) & II (G3) regions (2). The critical function of GTPases in several biological processes is illustrated by the involvement of these proteins, and factors stimulating their catalytic activity or mediating nucleotide exchange, in many physiological disorders including cancer.

The TRAFAC class of GTPases is subdivided into several superfamilies (1). Among them, the classical translation factor subgroup contains the well-known family of ubiquitous translation factors (EF-Tu/EF-1α, EF-G/EF-2, initiation and termination factors…) as well as three less well-characterized protein families named Bms1-like, HflX and OBG. Interestingly several of the latter factors were linked to ribosomes either through a direct role in translation or through their implication in ribosome biogenesis (e.g., references 3-8). The **D**evelopmentally **R**egulated **G**TP binding proteins (DRGs) belong to the OBG family of GTPases (which also contains the Obg/CgtA, YyaF/YchF, Nog1 and Ygr210 sub-branches). DRGs are strikingly conserved in archaea, fungi, plants

and animals. In addition to the GTPase domain, these proteins also contain a C-terminal TGS domain of unknown function. TGS domains are also found in other GTPases of the OBG family and is shared with threonyl-tRNA synthetases (ThrRSs) and guanosine polyphosphate phosphohydrolases/synthetases (SpoT/RelA) (the acronym TGS being derived from ThreRS, GTPase and SpoT,9). DRGs were first characterized by their abundant expression in mouse embryonic brain showing subsequent down-regulation in adult tissue (10,11). While archaea contains a single *drg* gene, two distinct DRG subtype, Drg1 and Drg2, are encoded by eukaryotic genomes (12). Some plants harbor three distinct genes, two of them code for nearly identical drg2 subtype proteins that are likely to result from a recent gene duplication event (13). Two-hybrid screens and co-immunoprecipitation experiments revealed that DRG GTPases interact with conserved partner proteins in yeast and human. Those were named **D**RG **F**amily **R**egulatory **P**rotein (DFRP). Dfrp1 (also known as Lerepo4 in human) binds specifically to Drg1 while Dfrp2 preferentially binds to Drg2 (14,15). Dfrp1 and Dfrp2 contain a C-terminal region of ~60 amino acids that was found to be required for binding to DRG and is named the dfrp domain (14). Else, Drfp1 and Drfp2 are highly divergent proteins, the former containing at its N-terminus two zinc fingers potentially mediating interactions with RNA while the latter contains a RWD domain that was identified in proteins interacting with the translational regulator Gcn1 (16). DFRP factor presence is important for the maintenance of normal levels of the cognate DRG proteins in human cells. Moreover, DRG-DFRP complexes were found to be localized in the cytoplasm of mammalian cells where the Drg1-Dfrp1 heterodimer was specifically found to associate with polysomes (17).

The yeast Drg1 homologue is named **R**ibosome **B**inding **G**TPase 1 (Rbg1) as it was found associated to ribosome (18,19). It associates with yeast Dfrp1, namely Tma46, which is also a ribosome associated protein (15,18). Consequently, yeast Drg2 was named Rbg2 (**R**ibosome **B**inding **G**TPase 2) even if, like its human counterpart, it fails to co-sediment with polysomes (15,17). Rbg2 associates with yeast Dfrp2, namely Gir2 (15). Consistent with the presence of a RWD domain, Gir2 was found to bind to Gcn1 (15,19). Yeast Rbg1 and Rbg2 are highly similar between themselves and with their human counterparts, Rbg1 sharing 66% identity and 80% similarity with human Drg1 and

Rbg2 59% identity and 80% similarity with human Drg2. The sequence conservation of DFRP factors between these two species is however much lower.

Although phylogenetic evidence and biochemical fraction studies have linked the DRG proteins to translation, differentiation and growth, the exact molecular function of these GTPases is as yet unknown. Early studies have suggested that mouse and human Drg1 interacts *in vitro* and *in vivo* with the oncogenic T-cell acute lymphoblastic leukemia (Tal1/Scl) protein, a basic helix-loop-helix (bHLH) transcription factor involved in cell growth and differentiation (20,21). It was also reported that overexpression of Drg1 increased rat embryonic fibroblast transformation induced by c-myc and *ras* overexpression, affecting both the onset and average size of foci formed (20). Drg2 was also reported to be downregulated in SV-40 transformed fibroblasts in comparison to normal fibroblasts (22). In other studies, mammalian Drg1 was also found to be a target for SUMOylation stimulated by the MEKK1 Map3 kinase (23) or shown to interact with the protein kinase MPSK1 (STK16) in a process requiring the N-terminal 65 residues of Drg1 (24). In yeast, filamentous invasion into agar matrices by *C. albicans* was attenuated by a Drg1 null mutation, concomitantly causing delayed lethality when the mutated organism was injected intravenously into mice. These phenotypes were suggested to result from the association of *C. albicans* Drg1 with Efg1 a bHLH transcription factor involved in repression of invasiveness (25). Many of these observations are difficult to reconcile with the conserved association of Drg1 factors to ribosomes. In yeast, deletion of *RBG1*, or *RBG2*, does not impair cell growth. Moreover, only very weak growth phenotypes resulting from double deletions of *RBG1* and *RBG2* could be detected using a sensitive competitive growth assay (26). An important step forward was made by the observation that a triple deletion mutant lacking *RBG1*, *RBG2* and the gene encoding the putative RNA helicase Slh1 exhibited a strong negative growth phenotype (15). Importantly, translation was found to be affected in this triple mutant, as evidenced by the presence of reduced levels of polysomes. Similar phenotypes were observed for other combinations of mutation inactivating simultaneously the Rbg1-Tma46, Rbg2-Gir2 and Slh1 functions, suggesting that these three entities mediate overlapping functions in translation (15).

To gain further insights into the function of Rbg1 and Tma46 and the mode of interaction of these two proteins, we decided to investigate the structure of this heterodimer. Only a few structures of GTPase of the OBG subfamily are currently known. This includes *B. subtilis* Obg (PDB id: 1LNZ) and human OLA1 of the YyaF/YchF subfamily (PDB id: 2OHF), the only structure available so far for a DRG subfamily member being the NMR solution of the C-terminal TGS of human Drg1. The crystal structure revealed the presence of novel domains in Rbg1 and uncovered the mode of interaction of Rbg1 with Tma46. Based on this information, *in vitro* and *in vivo* assays allowed us to dissect the role of the Rbg1-Tma46 domains and interactions in GTPase function and polysome recruitment.

Materials and methods

Plasmids

Plasmids were constructed using standard cloning strategies or by site-directed mutagenesis using the QuickChange strategy (Stratagene, La Jolla, CA, USA) with minor modifications to manufacturer's instructions. For expression constructs, an iterative trial and error process starting from plasmids encoding His6 tagged complete Rbg1 and full length Tma46 was used. Protein yields, subunit interaction and homogeneity were assessed by gel electrophoresis after purification on Ni-NTA. When necessary, mass spectrometry analyses, apparent fragment sizes and sequence comparisons were used in an attempt to define suitable domain borders. For protein production, expression plasmids were transformed into BL21-CodonPlus (DE3)-RIPL. In some instances, glycerol stocks of the transformed bacteria were stored at -80 ºC until use. Material obtained from plasmids encoding stable and well-expressed products were tested for crystallization. Yeast plasmids were constructed as described above and contained genes expressed under the control of their native promoters. All constructs were verified by sequencing.

All plasmids used in this study are listed in Supplementary Table 1 while oligonucleotides used to prepare these constructs are listed in Supplementary Table 2.

Purification of recombinant proteins

Protein expression was induced by growth in autoinduction media (Formedium, 27) plus kanamycin and chloramphenicol (50µg/ml and 24µg/ml, respectively) for 5 h at 37ºC followed by overnight growth at 20ºC. Cells were harvested by centrifugation and pellets kept frozen until further use. $Rbg1_{fl}$-$Tma46_{205-345}$ selenomethionine-substituted protein was obtained by using the autoinduction method (28,29). Small-scale protein production (100-200 mL) and purification were essentially performed as described earlier (30) except that BL21 CodonPlus was used for protein expression and buffer B (50 mM Tris HCl pH 7.4, 20 mM Imidazole, 300 mM NaCl, 2 mM ß-mercaptoethanol, 10% glycerol and 0.2% Igepal) for affinity purification on NiNTA. Proteins were eluted in buffer B containing 500 mM imidazole. For large scale preparations, pellets were thawed on ice and mixed with lysis buffer (350 mM NaCl, 20 mM Imidazole, 1 mM β-

Mercaptoethanol (β-MeOH), 20mM Tris pH 8.0 and a protease inhibitor cocktail tablet (Complete, EDTA-free, Roche) and sonicated. After centrifugation at 16,000 rpm (Sorvall RC-5C, SS34 rotor) at 4 ºC for 30 min, the lysate was filtered through a 0.45 µm sterile filter before loading onto a 5ml HisTrap FF Chelating column preloaded with 100 mM $NiSO_4$ and equilibrated in Buffer A (350 mM NaCl, 20 mM Imidazole, 1 mM β-Mercaptoethanol (β-MeOH), 20mM Tris pH 8.0). After washing the column with 10 bed volumes of Buffer A, proteins were eluted with a gradient of imidazole (20-500 mM) using an ÄKTA Purifier (GE Healthcare). Protein containing fractions were pooled and concentrated at 4 ºC to a final volume of 1-2 ml using Amicon Ultra centrifugal filter devices. The concentrate was directly loaded onto a pre-equilibrated size exclusion column (Sephadex 200 or 75 (16/60 or 26/60) columns) at 4 ºC and the protein eluted in buffer S (150 mM NaCl, 20 mM Tris pH 7.5 and 2 mM DTT) at rates of 1.0 ml/min using an ÄKTA Prime system. Purified proteins were then pooled and concentrated to 20-60 mg/ml by ultrafiltration at 4 ºC using the Amicon concentrator before flash-freezing in liquid nitrogen and storing at -80ºC. The purification protocol for the selenomethionine-substituted proteins was as above except that in the last step of the purification 5mM concentration of DTT was included to prevent selenomethionine oxidation.

Crystallization, data collection, structure refinement and analysis

Crystals of $Rbg1_{fl}$-$Tma46_{205-345}$ complex were obtained by the vapor diffusion method. The drops were setup at 4ºC with 1µl of 60 mg/ml of protein and 1 µl of reservoir solution (2.38 M Sodium formate, 0.2-0.5 M Sodium citrate pH 6.5). Three dimensional rectangular crystals with typical dimensions 0.3 mm x 0.05 mm x 0.02 mm grew in about 2 weeks**.** The X-ray diffraction data for the native and selenomethionine derivative were collected from single crystals at the beam line ID14-4 (31) at the European Synchrotron Radiation Facility at Grenoble, France using an ADSC Quantum Q315r CCD detector.

The data were indexed and integrated using MOSFLM (iMOSFLM) and scaled with SCALA in the CCP4 suite (32). Heavy atom site search and phasing were done using SHARP (33) and model building/tracing were done using ARP/wARP (34). Cycles of manual model building were performed with the program Coot (35). Waters were

introduced into the model using ARP/wARP program and validated with the electron density maps in Coot. The structures were refined with REFMAC (36) for isotropic refinement. TLS groups were defined and used for anisotropic refinement. This included 17 groups comprising of Rbg1 (chain A 2-45, 53-125/131-174/233-299, 175-232, 300-369; chain B 2-53, 54-91/98-125/133-174/233-299, 175-232, 300-369) and Tma46 (chain C 214-240, 241-267, 268-282, 302-313, 314-338; chain D 214-240, 241-267, 268-282, 320-336). Superimpositions between the structures were done using the SSM superpose function in Coot and analysis of the electrostatic surface potential was performed using APBS (37) in Pymol, also used for generating the structure figures (38).

GTP binding and hydrolysis assay

Thermal shift assay were performed using protein samples at around 0.05mM in buffer S with 5x Sypro Orange (Sigma), with or without GDP, GTP or GTPγS, in wells of MicroAmp 96-Well Fast Optical Reaction plate (Applied Biosystems). Fluorescence was measured from 20 to 85 ºC in increments of 1 ºC in a T7500 Fast Real-Time PCR system (Applied Biosystems). Results were analyzed using the GraphPad Prism 4 software (GraphPad Software Inc.).

GTP hydrolysis assays using Malachite green (39) were performed as follows: Solutions (5.72% w/v ammonium molybdate in 6N HCl, 0.08% w/v malachite green solution, 2.32% w/v polyvinyl alcohol) were prepared individually using reagents from Sigma Aldrich and stored at 4 ºC. For assays, MilliQ water: Malachite green: Polyvinyl alcohol: Ammonium molybdate were mixed in a ratio of 2:2:1:1 and incubated for 3 hours until they became yellow. 50 µl of protein samples (20µM) in filtered and degassed buffer (100mM KCl, 50mM Tris pH 7.5 and 5mM $MgCl_2$) were incubated with nucleotides (GTP or GDP, Sigma Aldrich) for 1 h at 37ºC in Microtest 96-Well Flat Bottom plates (Sarstedt). 200µl of Malachite green reagent was immediately added onto all the wells using a multipipette and, after 5 min incubation, the absorbance readings at 630 nm were measured in a Wallac Victor2 1420 Multilabel Counter. A phosphate standard prepared from KH_2PO_4 and blank with no protein were included on the plate. The latter background was subtracted from the protein sample readings. Data were fitted

to Michaelis-Menten equation using non-linear regression in GraphPad Prism 4 to determine the kinetics.

RNA homopolymer binding assay

For non-specific RNA binding studies, proteins were incubated with poly(U)-agarose (Sigma). Approximately 10 µl of a 50 % bead suspension mixed with 0.2-0.4 mg of protein were nutated in a 4 ºC room for 30 min in 50µl of reaction buffer (100mM NaCl, 10mM Hepes pH 7.4, 2mM $MgCl_2$, 0.1% Triton X-100, 3mM DTT and 0.1 mg/ml BSA). The beads were washed six times in 500 µl of wash buffer (reaction buffer without BSA) before SDS elution of bound proteins in SDS sample buffer, fractionation in a 10% SDS-polyacrylamide gel and visualization by Coomassie staining. BSA and poly(U)-agarose beads treated with 1 mg/ml of RNAse A prior to protein addition were used as negative control. For competition assays, poly(U) (Sigma) at concentrations of 0, 0.1 and 1 mg/ml was pre-incubated with proteins in reaction buffer for 20 min at 4 ºC prior to beads addition. Heparin at concentrations of 0.5, 1, 2 and 4 mg/ml were added to the reaction mixture prior to beads addition to analyze the strength of binding.

Yeast strains and growth assays

Yeast strains are all derived from BMA64 strains (40) and are listed in Supplementary Table 3. Strains containing a single disrupted and epitope-tagged gene were obtained by transformation with TAP-tag (41) and HISMX6 (42) modules carrying short flanking sequences homologous to the targeted gene. Primer sets that were used for that purpose are described in Supplementary Table 2. Transformants were checked for correct integration by PCRs. Plasmids were introduced into yeast strain using the standard LiOAc transformation method (43).

For growth assays, yeast cultures were grown to saturation in selective liquid media. The cultures were then diluted in water to an optical density at 600 nm (OD_{600}) of 0.1. Three microliters of these cultures and 10-fold serial dilutions were spotted onto agar plates containing complete synthetic media minus leucine. Plates were incubated at 37℃ and 30℃ for 3 or 4 days, and cell growth was determined by visual inspection and documented by photography.

Western blot analyses

Proteins from immunoprecipitation experiments, or total yeast extract (44), were fractionated by SDS-PAGE and subsequently transferred to nitrocellulose membrane. TAP tagged proteins were detected as described earlier (45). HA- tagged proteins were detected using mouse anti HA monoclonal antibody (Covance MMS-101P) and a secondary goat-anti-mouse IgG antibody (Jackson 115-035-068). As loading control Stm1 was detected by a polyclonal anti-Stm1 antibody and a secondary goat-anti-rabbit IgG antibody (Pierce 31460). Chemiluminescence was recorded with a LAS4000 device (GE Healthcare).

Immunoprecipitation of epitope tagged-proteins

Logarithmically growing yeast cells in selective medium at 30°C were resuspended in lysis buffer containing 10mM Tris HCl pH 7.5, 150mM NaCl, 5mM $MgCl_2$, 1mM DTT and protease inhibitors. Cells were broken by mixing with glass beads. The cell extract was obtained by two consecutive centrifugations, the first 20 min at 14,000 g and the second 10 min at 14,000 g. Glycerol was added to a final concentration of 10%. 30 mg/ml of total proteins were incubated with IgG Sepharose Beads (GE Healthcare) or IgG coupled to Dynabeads on a rotor at 4℃ for 2 hours. Beads were pelleted and washed extensively with IPP150 buffer (10mM Tris HCl pH 7.5, 150mM NaCl, 5mM $MgCl_2$). Proteins bound to the beads were eluted with SDS-PAGE sample buffer by boiling for 5 minutes.

Polysome analyses

Polysomes were analyzed essentially as described previously (15).

Results

Structure determination of the yeast Rbg1-Tma46$_{(205-345)}$ complex

To obtain recombinant Rbg1 and Tma46, or truncated derivatives thereof, for structural and functional analyses, we constructed artificial operons encoding various versions of 6His-tagged Rbg1 followed by Tma46. Initially, full-length Rbg1 and Tma46 were used. The two proteins co-purified indicating that no yeast specific factor or compound was necessary to allow their interaction. However, low complex yields and truncated forms of Tma46 were observed. Mass spectrometry analyses and estimation of apparent molecular weights, together with sequence analyses delineating borders of conserved domains, provided rough estimates of the missing regions. After several iterative cycles of construct optimization, a plasmid expressing efficiently and without apparent degradation 6His-tagged full-length (fl) Rbg1 together with the C-terminal region of Tma46 (amino acids 205-345) encompassing the DFRP region was obtained.

The X-ray crystal structure of proteins obtained with the latter construct was solved to 2.67Å resolution by the SIRAS method using a selenomethionine-substituted protein. The final model of the Rbg1-Tma46$_{205-345}$ complex (R factor 19.7%, R_{free} 22.2%) includes two molecules of the complex in the asymmetric unit although the complex behaves as a heterodimer by size exclusion chromatography. The structure is described with molecule A and B (Rbg1) interacting with molecule C and D (Tma46), respectively. The data collection and refinement statistics are as given in Table 1. Rbg1 was modeled from 2 to 369 in both molecule A and B, but some of the loops, in particular those comprising the G-motifs, had poor electron density due to their flexibility. 77 water molecules in the first solvation shell were included.

The Rbg1 structure shows a domain organization that includes the well-conserved G-domain (G1+G2+G3 = 64-169; G4+G5 = 245-293), an N-terminal helix-turn-helix (HTH) subdomain (1-44) which lies adjacent to another subdomain formed by a 65 residue long insertion (176-240) between G3 and G4 of the G-domain and the TGS domain at the C-terminus (294-369) (Figure 1).

The G-domain of Rbg1 is highly similar to the well-conserved GTP binding domain of other GTPases, in particular proteins Obg, YchF, FeoB, HflX, Ras-related proteins or Era, which all belong to the TRAFAC class of GTPases. The G-domain

contains five α-helices (α3, α4, α5, α8 and α9) and six β-strands (β1, β2, β3, β4, β10 and β11). The G1 motif (P-loop/Walker A) is located on the loop connecting β1 to α3, G2 (Switch I) in between α3 and β2, G3 (Switch II/Walker B) at the end of β3 strand, G4 in β10 and G5 motif in β11 (Figure 1).

Previous reports have shown that DRG factors contain about 65 amino acids inserted between the G3 and G4 motifs of the G-domain that are not found in other Obg family members and had no sequence homology to known domains (19). This region (residues 176-240) folds as an independent domain emerging from the G domain. We named this new domain of Rbg1 the S5D2L domain as database searches indicate that its topology is related to the "Ribosomal protein S5 domain 2-like" (S5D2L) superfamily despite the absence of significant sequence similarity. The latter superfamily has 13 members (as according to Pfam, CATH and SCOP) and structural alignment shows that whereas the other members have a βββαβα fold, S5D2L domain has ββαβα fold lacking the first β strand (Supplementary Figure 1). The bacterial 30S ribosomal S5 subunit protein C-terminal domain is structurally the most similar to S5D2L domain, aligning with an RMSD of 2.2 Å over 51 residues although the sequence identity was very low (12%). Interestingly, however, Rbg1 Gly189 and Arg207, mutations of which cause ribosomal ambiguity (RAM) (46), are fully conserved in S5 subunit protein. Equivalent residues are also present in EF-G domain IV. The nature of these residues is however not universally conserved, as they are not found in GHMP kinase family members. A short parallel beta sheet formed by β5 and β9 of the S5D2L domain tether this structure on the back of the G domain between the segments containing the G3 and G4 motifs.

The two amphipathic helices, α1 and α2 (2-44), comprises a previously unnoticed Helix-Turn-Helix (HTH) subdomain at the N-terminal of Rbg1 which lies adjacent to the S5D2L domain. The HTH is stabilized mainly by a hydrophobic zipper between the two helices, five leucines positioned 3-4 residues apart in the longer helix α2 contributing to the zipper while the other side of the helix α2 makes both hydrophilic and hydrophobic interactions with the S5D2L helices α6 and α7. Additionally, the interface between α1 and α2 is also stabilized by hydrogen bonds between Glu12 (α1) - Arg39 (α2) and Glu14 (α1) and His27 (α2).

The C-terminal TGS domain (294-369) has predominantly beta sheet structure with five beta strands (β12-β16) and two helices, α10 and α11 (Figure 1). Superimposition of the Rbg1 TGS domain with the previously reported NMR solution structure of human Drg1 TGS domain (PDB id: 2EKI) gave an RMSD of 1.31 Å over 76 Cα atoms. The TGS domain was found to be structurally very similar to threonyl tRNA synthetase, YchF and hOLA1 TGS domains and ubiquitin.

Both the TGS and S5D2L-HTH domains lay on the distal part of the GTP binding pocket (Figures 1 and 2). Electrostatic surface potential analysis shows an extensive positively charged surface formed partly by the TGS, HTH-S5D2L and part of the G-domain opposite to the GTP binding site (Figure 2).

Tma46 structure and interaction with Rbg1

Tma46 fragment present in the structure shows a non-globular type of fold predominantly formed by alpha helices interconnected by coils. The residues of $Tma46_{205-345}$ in the complex with Rbg1 that we could model into the electron density are located as mainly four helices (numbered in this study as α1- α4) and a short beta strand which forms beta sheet structure with adjacent beta strands β2, β3, β1, β4, β10 and β11 from the G-domain of Rbg1. Tma46 helices envelops Rbg1 forming an extended and extensive interface (buried surface area in the interface about 2978 $Å^2$) contacting the G-domain (helices α1 and α2) and the TGS domain (helices α3, α4 and the beta sheet) with no contacts with the remaining domains of Rbg1, namely S5D2L and HTH (Figure 1 and 2). Based on weak sequence similarities between Dfrp1 and Dfrp2, it was earlier proposed that the region corresponding to residues 280 to 332 of Tma46 constituted a DFRP domain responsible for interaction with Rbg1 (14). Based on two-hybrid screen results, it was also previously suggested that Tma46 residues 254-296 would constitute the Rbg1 binding site (19). Our complex structure demonstrate that the region of Tma46 contacting Rbg1 is larger than these earlier estimates and encompass residues 216-279 and 302-338. This observation suggests further that the DFRP region defined earlier is only a fraction of the biologically relevant unit involved in DFRP-DRG protein interaction and that the region of Gir2/Dfrp2 involved in contacting Rbg2 is also larger.

One of the most prominent interacting surfaces between Rbg1 and Tma46 comprises α8 helix of Rbg1 (between G4 and G5 motifs), which inserts between helices α1 and α2 of Tma46 themselves interacting in turn also with Rbg1 α5 and α9 respectively. Helix α2 is longest in the fragment of Tma46 solved, and extends from the Rbg1 G-domain to the TGS domain that it contacts with its C-terminal end. Interestingly, the interface between Tma46 α2 and Rbg1 α8/α9 consists mainly of aromatic ring containing residues, which form a pi stacking interaction. Specifically, Tyr264 (α8) and Trp278 (α9) of Rbg1 form stacking interactions with Phe246, Trp249 and His253 of α2 of Tma46 (Supplementary Figure 2).

Interestingly, we noticed that while the G-domain of chain A and chain B adopt the same overall fold, (RMSD between the chains over the backbone Cα atoms is 1.38 Å) differences were observed especially in the loops containing the five G-motifs. By superimposing human OLA1 structure bound to ATP onto the G-domain of the two chains, it was seen that the P-loop in chain A adopted a closed conformation where entry of GTP could be difficult whereas that of chain B had an open conformation. The loops containing the G2 and G3 motifs were also shifted although the electron density in this area was not complete. We cannot however rule out that the observed conformational difference in this area might be due to crystal packing contacts.

Dissecting Tma46 interaction with Rbg1

These structural data allowed us to analyze in detail how Tma46 recognize Rbg1. As this involves a fragment of Tma46 that does not fold as a globular domain, but rather as a string of independent structural elements that meander on the surface of Rbg1, it is likely that Tma46 is intrinsically unfolded and only adopt the observed conformation upon binding to Rbg1. Interestingly, this is likely to also apply to Gir2 which was shown to be intrinsically unstructured (47).

We constructed stepwise deletion of a HA-tagged version of the *TMA46* gene inserted in a yeast vector, removing defined structural elements that interact with Rbg1. Four mutants removing successively the Tma46 ß strand, helix α4, helix α3, helix α2 from the C-terminus and two mutants removing successively helix α1 and helix α2 from the N-terminus were built. We first tested whether these mutants were functional by

assaying their ability to complement the triple mutant Δ*gir2*Δ*tma46*Δ*slh1* for its slow growth phenotype, a feature exacerbated at 37°C (15). We controlled by western blotting that the mutant proteins were expressed and accumulated to normal levels (Figure 3). A control plasmid encoding a complete Tma46 restored a wild type phenotype demonstrating that the presence of the tag does not impact on its function. Deletion of the Tma46 ß strand, alone or in combination with helix α4 does not impair Tma46 function. Removing the ß strand with helices α4 and α3 partially disrupts Tma46 activity while removing the region extending from the C-terminus and including helix α2 inactivate Tma46 (Figure 3A). Tma46 mutant levels were normal (Figure 3B) demonstrating that the partial activity did not result from protein instability but rather from inactivity. Reciprocally, deletion of helix α1 had limited effects on Tma46 level or function whereas deletion of helices α1 and α2 destabilized the protein resulting concomitantly in a poor complementation (Figure 3A and B).

We next assessed whether these mutations of Tma46 affect interaction with Rbg1 *in vivo*. For this purpose, plasmids encoding the mutant were introduced in a Δ*tma46* strain carrying a TAP-tagged Rbg1 allele. Extracts prepared from transformants were incubated with IgG beads to precipitate TAP-tagged Rbg1 and associated factors. Presence of Tma46, and as a control of Rbg1-TAP, in the input and immunoprecipitated fractions (pellets) was assessed by western blotting (Figure 3 C and D). This analysis demonstrated that deletions of the Tma46 ß strand, of the ß strand with helix α4, or of helix α1 do not affect interaction with Rbg1. In contrast, Tma46 proteins lacking larger fractions of the interaction region (ß strand with helices α4 and α3, ß strand with helices α4-2, or helices α1 and α2) do not co-precipitate with Rbg1 indicative of strongly reduced or absent interactions. Interestingly, the results of the interaction analysis and phenotypic interaction correlate well, indicating that interaction is synonym of function. A single exception was observed with the mutant lacking Tma46 ß strand and helices α4 and α3 that is negative for interaction but that partly complements the Tma46 inactivation *in vivo*, being functional at 30°C but not at 37°C (Figure 3A). This result is likely to result from a limited residual interaction of this truncated Tma46 with Rbg1 that did not withstand co-immunoprecipitation.

Altogether, our data demonstrate that several elements are implicated in ensuring interaction of Tma46 with Rbg1. The presence of superfluous elements, demonstrated by the lack of functional phenotype and effect on interaction of several mutants, may ensure an extremely tight binding. At this stage, all elements of Tma46 mediating interaction with Rbg1, except helix α2, appear not to be essential for function. It is likely that helix α2 is important given that its interaction with Rbg1 occur close to the GTPase center (Figure 1 and 2).

GTP binding and hydrolytic activity of Rbg1 is modulated by interaction with Tma46

DRG proteins have been shown to bind GTP and GDP (11,48). Moreover, Arabidopsis DRGs have been reported to hydrolyze GTP into GDP *in vitro* without the help of GAPs or GEFs, unlike Ras-like proteins (13). To investigate the effect of complex formation on GTP binding by Rbg1, we performed thermal shift assay in the presence or absence of GTP, GDP or GTPγS for Rbg1 and Rbg1 with the C-terminal part of Tma46 (Figure 4). Nucleotide binding of the complex was evidenced by an increased melting temperature in the presence of the GDP, GTP or non-hydrolysable GTPγS. Moreover, comparison of free Rbg1 and the Rbg1-Tma46 complex revealed that the increased stability detected in the presence of GTPγS was specific for the complex, suggesting that complex formation favors nucleotide binding.

We next analyzed the GTP hydrolytic activity of free-, and Tma46-bound-, Rbg1. The $Rbg1_{fl}$-$Tma46_{205-345}$ complex shows a weak GTP hydrolysis activity (Fig. 4, Table 2), with parameters similar to those reported for other Obg family of proteins (49,50) and atDrg1/atDrg2a (13). A catalytic site mutant S79N (G1 motif changed from GFPSVGKS to GFPSVGK**N,** which was shown to be inactive *in vivo* (15) did not show significant hydrolytic activity above the detection limit of the assay (Figure 4B) confirming the specificity of the reaction. Additionally, the recombinant human homolog, $Drg1_{fl}$-$Lerepo4_{220-396}$ complex was about twice as active as the cognate yeast complex (Table 2). For comparison, Rbg1 and Drg1 on their own showed significantly reduced GTPase activity (Fig. 4). The increased activity of the complex over free Rbg1/Drg1 could be attributed mostly to the stabilizing function of Tma46/Lerepo4. This is evidenced by a

decrease in activity of Rbg1 in complex with a $Tma46_{dfrp}$ mutant lacking the helix α1 (Table 2). Furthermore, mutants of residues of Tma46 in the interface between Rbg1 (helix α8) and Tma46 (helix α2) such as Ile241, Phe246, Trp249 and Lys250, the latter three of which are involved in pi stacking, also show reduced activity (Table 2).

We also tested whether the presence of RNA affected the GTPase activity by incubating $Rbg1_{fl}$-$Tma46_{205-345}$ with poly(U) for 20 min and then performing the GTPase assay. The kinetic parameters were similar to the native protein suggesting that RNA binding does not influence GTPase activity *in vitro*.

Role(s) of the different Rbg1 domains

The structure of the Rbg1-Tma46 heterodimer indicates that Rbg1 comprises four domains. The GTPase domain is the largest of them forming a platform surrounded by protuberances corresponding to the HTH, S5D2L and TGS domains. Previous analyses indicated that the GTPase and TGS domains of Rbg1 were essential for its function (15). However, careful examination of the structural data now available indicates that the TGS deletion used probably extended into the GTPase domain. This situation was probably responsible for the instability of the mutant protein (15). The HTH and S5D2L had not been tested, as sequence analyses had not identified them unambiguously. We therefore constructed precise deletions of the HTH, S5D2L or TGS domains in a functional HA-tagged version of the *RBG1* gene inserted in a yeast vector. As deletion of the S5D2L removes an internal part of the protein, two mutants were built incorporating one or two glycine residues at the deletion point to allow sufficient flexibility and length of the polypeptide backbone to allow folding.

These mutants were first assayed for their function through their ability to complement the slow and temperature sensitive growth of a Δ*rbg1*Δ*rbg2*Δ*slh1* strain (Figure 5A). Deletion of the HTH and TGS domain inactivated Rbg1. Interestingly, removal of the S5D2L complemented efficiently the mutant strain at 30°C but was unable to do so at 37°C (Figure 5A) even though the protein was stably expressed at both temperature (data not shown). This result demonstrates that deletion of the S5D2L domain did not disrupt the protein but that the presence of the S5D2L domain is essential for the full activity of Rbg1. Western blot analysis confirmed that all protein were

expressed at (or near) wild type levels indicating that the HTH, S5D2L or TGS domains are required for Rbg1 activity rather than stability (Figure 5B).

We next analyzed the roles of the different Rbg1 domain in mediating interaction with Tma46. Plasmids encoding full-length or truncated versions of Rbg1 together with regions encompassing the C-terminus of Tma46 (residues 154-345 or 205-345) were constructed. After expression in *E. coli*, the recombinant proteins were purified on Ni-NTA thanks to the presence of a 6His tag inserted at the N-terminus of the Rbg1 fragment and protein eluates were analyzed by denaturing gel electrophoresis and Coomassie staining (Figure 5C). This analysis demonstrated that an inactive GTPase or a form of Rbg1 lacking the S5D2L still interacted with Tma46. In contrast, deletion of the TGS prevented Tma46-Rbg1 interaction (even though low residual binding was occasionally detected). Interestingly, the TGS domain by itself was sufficient to interact with Tma46. Unfortunately, the effect of deleting the HTH domain could not be addressed in this assay because the corresponding protein was poorly expressed in *E. coli*.

To analyze the implication of the different Rbg1 domains in interaction with Tma46 *in vivo*, we introduced the shuttle plasmids encoding HA-tagged truncated versions of Rbg1 in a yeast Δ*rbg1* strain expressing Tma46-TAP. Co-precipitation of the two proteins was assayed as described above (Figure 3). The results of this experiment demonstrated that deletion of the HTH or S5D2L had no impact on the capacity of Rbg1 to bind Tma46 while no interaction was detected after removal of the TGS domain (Figure 5D).

Altogether, consistent with the structure of the complex, these results demonstrate that the HTH and S5D2L domains do not contribute to Tma46 binding. These data reveal, however, a critical role for the TGS domain in the association of Tma46 with Rbg1.

We also assayed the GTPase activity of deletion mutants of Rbg1 lacking the TGS or S5D2L domains co-expressed with the C-terminal region of Tma46 ($Tma46_{205-345}$). All mutants displayed a significant GTP hydrolytic activity albeit with reduced rate compared to the wild type protein (Table 2). This result indicates that these two peripheral domains of Rbg1 are not essential for GTPase activity although they may modulate catalysis.

Assay for RNA homopolymer binding

To test the role of the various domains of Rbg1 and of Tma46 in RNA binding, we performed a RNA homopolymer binding assay similar to the one previously described for *Xenopus laevis* Drg1 and Drg2 (51). Rbg1 associated with $Tma46_{205-345}$, but not BSA used as a negative control, bound to poly(U) (Fig. 6). Pre-treatment of beads with RNase completely eliminates binding confirming that association with beads occurs through RNA (Figure 6). Consistently, binding was competed by increasing concentrations of free poly(U), and to a lower extent by heparin indicating a strong affinity for poly(U). An active GTPase was not required for binding as observed with the S79N Rbg1 substitution mutant (data not shown). Consistent with previous results obtained for Xenopus Drgs (51), Rbg1 alone was able to bind to RNA, although with an apparent lower affinity (Fig. 6). This property extended to the human homologue.

Several proteins with S5D2L domains were shown to bind RNA, yet the S5D2L domain of Rbg1 did not bind to poly(U) whereas the complex Rbg1-Tma46 complex lacking the S5D2L still bound RNA efficiently (Fig. 6C). Interestingly, Rbg1 lacking the TGS domain and the TGS domain alone bound to $Tma46_{205-345}$ both interacted with RNA. These observations suggest that several domains of Rbg1, including the TGS domain interact with RNA (Fig. 6C).

Tma46 recruits Rbg1 in polysomes

Like their human homologues, Rbg1 and Tma46 have been shown to associate with polysomes (15,17,19). Thus, we next assessed the ability of Rbg1 deletion mutants to associate with polysomes. Extracts prepared from strains expressing wild type TAP tagged Tma46 and the various mutant forms of Rbg1 tagged with an HA epitope were layered on sucrose gradients. After centrifugation, fractions of the gradients were collected while monitoring RNA absorbance (Figure 7). Presence of Tma46 and wild type or mutant forms of Rbg1 in the various fractions was monitored by western blotting.

As reported earlier (15), wild type Rbg1 and Tma46 co-sediment in polysomes (Figure 7A). A similar distribution was observed when the HTH or S5D2L domains were deleted, albeit increased free Rbg1 was detected at the top of the gradient in the latter case (Figure 7C and D). Interestingly, removal of the TGS domain from Rbg1 resulted in complete segregation of the two proteins with Tma46 being localized in polysomes while

Rbg1 remained at the top of the gradient (Fig. 7B). Taken together with the observation that the TGS domain is sufficient and necessary to mediate interaction between Rbg1 and Tma46 (Figure 5), this observation is consistent with Tma46 mediating the recruitment of Rbg1 in polysomes.

Discussion

DRG proteins are extremely well conserved factors present in eukaryotes and archaea that belong to the Obg/Drg GTPase subfamily (1). Characterized members of the latter group have all been implicated in ribosome genesis or function. Intriguingly, the two highly related DRG GTPases found in eukaryotes associate with rather distantly related DFRP co-factors, but still mediate partly overlapping functions.

To gain insights into the mode of action of these puzzling proteins, we determined the crystal structure of one of the yeast DRG, namely Rbg1, in complex with a fragment of its partner Tma46, and analyzed the functional roles of newly uncovered structural elements. Our results provide unexpected insights into the organization of DRG GTPases and into their mode of interaction with DFRP partners, including the impact of the latter on GTPase activity. Moreover, our data demonstrate a key role of Tma46 in mediating Rbg1 recruitment to polysomes. Sequence analyses suggest that a distantly related mechanism is exploited by the second eukaryotic DRG/DFRP pair (see below).

Sequence alignments had revealed the presence of a GTPase and TGS domains in DRG proteins, including Rbg1. Moreover, the presence of an insertion implanted between conserved GTPase signature elements had been noticed (19). Our structure demonstrates that this insertion folds as an independent domain that protuberates on the surface of the Rbg1 GTPase module. This inserted domain, S5D2L, adopts a fold related to domain 2 of the bacterial S5 protein. This ribosomal protein paves the entrance of the mRNA channel in the *E. coli* ribosome and is, interestingly, the target of mutations affecting translational fidelity (52,53). The presence of a related domain in a GTPase linked to ribosomal activity may thus be of functional significance. The structure of Rbg1 reveals in addition the presence of a helix-turn-helix (HTH) at the extreme N-terminus of the protein, whose presence had not been foreseen by sequence analyses. The HTH domain packs against the S5D2L while the TGS domain is distantly located in space and thus do not contact these

two units. Overall, the Rbg1 protein, and related DRG factors, appear to be a multidomain GTPase with the GTPase core forming a platform on which are grafted three structural domains forming two independent protuberances (HTH and S5D2L on one side, and TGS on the other side). Mutational analysis indicates that the GTPase activity is absolutely required for Rbg1 function. Deletion of the HTH or TGS domain also abrogates Rbg1 activity, while, interestingly, removal of the S5D2L results in a conditionally active protein that is unable to function at high temperature. Thus, all domains of Rbg1 are required for its full activity. The S5D2L domain is not absolutely essential for function and may either stimulate protein activity under extreme conditions or be substituted by other factors in less demanding situations.

In the crystal structure, the C-terminal fragment of Tma46 adopts an extended conformation (Figure 1A) made essentially of successive α-helices interconnected by loops that meanders on the surface of the GTPase and TGS domains of Rbg1. Interestingly, part of Tma46 embraces an area of Rbg1 that is close to the GTP binding site. Consistently, we observe that the presence of Tma46 affect GTP binding and hydrolysis by Rbg1. But, at this stage, it is not possible to conclude whether this results from a local effect or from a more general impact on Rbg1 stability. Deletions analyses indicate that the length of the segment of Tma46 interacting with Rbg1 can be reduced from either extremities without abrogating binding or abolishing function. Unexpectedly, most structural elements of this region of Tma46 are individually not essential, even if they contribute to binding and function. The central α2 helix remains the only structural element potentially absolutely required for function. However, no definitive statement regarding the functional essentiality of helix α2 can be made yet as a construct removing helix α2 while maintaining a configuration of the remaining segments of Tma46 allowing them to interact productively with Rbg1 may be exquisitely complex to design.

The region of Tma46 interacting with Rbg1 shows only limited conservation and is larger than the segment previously identified as the Dfrp domain based on protein alignment. This part of Tma46 is unlikely to adopt a globular fold on its own, and thus probably corresponds to an intrinsically unstructured polypeptide. Accordingly, the Gir2 protein, that contains an equivalent region, has been shown to be intrinsically unstructured (47). Our results demonstrate that Tma46 contains structural elements superfluous for

interaction with Rbg1 and function. Taken together with the lack of globular structure of this region and its extended conformation, it is tempting to propose that Tma46 evolved by the successive additions and extensions that increased its ability to interact with the Rbg1 surface. Such an evolutionary model is easy to imagine and fully consistent with the lack of obvious Tma46 or Gir2 homologs in archaea despite the presence of DRG homologues. Moreover, this framework also provide for an explanation for the poor sequence conservation of the Gir2 and Tma46 regions mediating interaction with Rbg2 and Rbg1 in yeast, and in the homologous proteins from other species, despite the extraordinary conservation of DRG factors. Indeed, any substitution arising in Dfrp protein still allowing efficient interaction with DRG factors will be accommodated, even if this results in a slightly altered relative structural arrangement. Overall, these constraints will result in an asymmetric rate of evolution of the two partners, with DRG factors changing slowly over time and DFRP proteins exploring rapidly an extensive sequence space. The availability of the Rbg1-Tma46$_{205\text{-}345}$ structure might have provided explanation on why Rbg1 interacts with Tma46 *in vivo* while Rbg2 interacts with Gir2 despite the strong similarity between the two yeast DRG factors (15,17). However, mutagenesis of candidate residues failed to identify key amino acid interaction networks essential to establish this specificity (data not shown). This suggests that specificity is based on a large set of interactions rather than a few key amino acids. It is noteworthy also that the specific interaction of one DRG factor with a DFRP partner is not absolute, as cross-interaction can be detected in artificial conditions (19).

Contrasting with our observation that deletion of the Tma46 ß strand and helices α3 and α4 only reduces Tma46-Rbg1 interaction *in vivo* resulting in temperature sensitive function, deletion of the Rbg1 TGS domain abrogates the Rbg1-Tma46 interaction, both *in vivo* and in assays based on recombinant proteins. This observation is surprising because the region of Tma46 that interacts with the Rbg1 TGS domain is composed of helices α3, α4, and of the ß strand. The latter result may indicate that the TGS domain contributes to Tma46 binding both by providing an extensive surface of contact for Tma46 and by maintaining the GTPase domain in a conformation favorable for interaction. □□□□□□□assays using recombinant factors confirm that Tma46 interacts efficiently with the isolated TGS while it binds inefficiently, at best, with a truncated

form of Rbg1 lacking the TGS domain. These data support the idea that the stability of the Rbg1-Tma46 interaction resides for a major fraction in the area involving the TGS domain and/or that the TGS domain constitutes a primary nucleating center for the formation of this heterodimer. This demonstrate that the Rbg1 TGS domain has a critical role in mediating protein interaction and it will be of interest to test whether this property extends to TGS domains present in other proteins.

Interestingly, deletion of the TGS domain *in vivo* resulted in the fractionation of Tma46 and the mutant Rbg1 in distinct regions in polysome gradients: while Tma46 remained associated with polysomes, Rbg1 was released and found as a free factor in the lighter fractions of the gradient. This is a specific effect of the TGS deletion, as similar distributions were not observed with removal of the S5D2L or HTH domains of Rbg1. This suggests that Tma46 associates with polysomes, possibly through an interaction of its two Zn fingers with mRNAs and/or ribosomal RNAs, thereby recruiting Rbg1 in these assemblies. These observations strengthen the role of Rbg1 in translation and suggest that its GTPase would be able to mediate a yet-to-be-defined action on ribosomes after its recruitment. Interestingly, a variant of this model may apply to Rbg2 as well. Indeed, in specific conditions, an interaction of the N-terminal RWD domain of Gir2 with ribosome bound Gcn1 may provide a means to recruit Rbg2 to translating ribosome in a manner similar to the recruitment of Gcn2 (16). Further structural and biochemical work will be required to test the biological relevance of this hypothesis.

Supplementary data

Supplementary Data are available at NAR online: Supplementary Tables 1-3, Supplementary Figures 1-2, and Supplementary References (15,40,54).

Funding

MEG received postdoctoral fellowship from the Ministerio de Ciencia e Innovación and Conselleria d'Educació, Generalitat Valenciana. This work was supported by the Ministerio de Ciencia e Innovación (SAF2008-04048-E, SAF2009-10667) to JB, Conselleria de Sanitat, Generalitat Valenciana AP-001/10 to JB, CSIC (200820I020) to JB, by the European Union 6th Framework programs 3D Repertoire (contract LSHG-CT-2005-512028) to JB and BS, by the CERBM-IGBMC to BS, the CNRS to BS and the Ligue Contre le Cancer (Equipe Labellisée 2011) to BS.

Acknowledgements

The authors thank J. Létoquart and V. Henriot for preparing and testing numerous constructs that allowed recombinant protein production, M. Correa for technical support, M. Argentini and D. Cornu (SICaPS, IMAGIF platform, Gif/Yvette, France) for mass spectrometry analysis and the IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire) for assistance. We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities and we would like to thank staff at beamline ID14-4 for technical support and assistance.

References:

1. Leipe, D.D., Wolf, Y.I., Koonin, E.V. and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol*, **317**, 41-72.
2. Paduch, M., Jelen, F. and Otlewski, J. (2001) Structure of small G proteins and their regulators. *Acta Biochim Pol*, **48**, 829-850.
3. Gelperin, D., Horton, L., Beckman, J., Hensold, J. and Lemmon, S.K. (2001) Bms1p, a novel GTP-binding protein, and the related Tsr1p are required for distinct steps of 40S ribosome biogenesis in yeast. *RNA*, **7**, 1268-1283.
4. Wegierski, T., Billy, E., Nasr, F. and Filipowicz, W. (2001) Bms1p, a G-domain-containing protein, associates with Rcl1p and is required for 18S rRNA biogenesis in yeast. *RNA*, **7**, 1254-1267.
5. Polkinghorne, A., Ziegler, U., Gonzalez-Hernandez, Y., Pospischil, A., Timms, P. and Vaughan, L. (2008) Chlamydophila pneumoniae HflX belongs to an uncharacterized family of conserved GTPases and associates with the Escherichia coli 50S large ribosomal subunit. *Microbiology*, **154**, 3537-3546.
6. Kallstrom, G., Hedges, J. and Johnson, A. (2003) The putative GTPases Nog1p and Lsg1p are required for 60S ribosomal subunit biogenesis and are localized to the nucleus and cytoplasm, respectively. *Mol Cell Biol*, **23**, 4344-4355.
7. Jensen, B.C., Wang, Q., Kifer, C.T. and Parsons, M. (2003) The NOG1 GTP-binding protein is required for biogenesis of the 60 S ribosomal subunit. *J Biol Chem*, **278**, 32204-32211.
8. Scott, J.M., Ju, J., Mitchell, T. and Haldenwang, W.G. (2000) The Bacillus subtilis GTP binding protein obg and regulators of the sigma(B) stress response transcription factor cofractionate with ribosomes. *J Bacteriol*, **182**, 2771-2777.
9. Wolf, Y.I., Aravind, L., Grishin, N.V. and Koonin, E.V. (1999) Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res*, **9**, 689-710.
10. Sazuka, T., Kinoshita, M., Tomooka, Y., Ikawa, Y., Noda, M. and Kumar, S. (1992) Expression of DRG during murine embryonic development. *Biochem Biophys Res Commun*, **189**, 371-377.
11. Sazuka, T., Tomooka, Y., Ikawa, Y., Noda, M. and Kumar, S. (1992) DRG: a novel developmentally regulated GTP-binding protein. *Biochem Biophys Res Commun*, **189**, 363-370.
12. Li, B. and Trueb, B. (2000) DRG represents a family of two closely related GTP-binding proteins. *Biochim Biophys Acta*, **1491**, 196-204.
13. O'Connell, A., Robin, G., Kobe, B. and Botella, J.R. (2009) Biochemical characterization of Arabidopsis developmentally regulated G-proteins (DRGs). *Protein Expr Purif*, **67**, 88-95.
14. Ishikawa, K., Azuma, S., Ikawa, S., Semba, K. and Inoue, J. (2005) Identification of DRG family regulatory proteins (DFRPs): specific regulation of DRG1 and DRG2. *Genes Cells*, **10**, 139-150.
15. Daugeron, M.C., Prouteau, M., Lacroute, F. and Seraphin, B. (2011) The highly conserved eukaryotic DRG factors are required for efficient translation in a manner redundant with the putative RNA helicase Slh1. *Nucleic Acids Res*, **39**, 2221-2233.

16. Nameki, N., Yoneyama, M., Koshiba, S., Tochio, N., Inoue, M., Seki, E., Matsuda, T., Tomo, Y., Harada, T., Saito, K. *et al.* (2004) Solution structure of the RWD domain of the mouse GCN2 protein. *Protein Sci*, **13**, 2089-2100.
17. Ishikawa, K., Akiyama, T., Ito, K., Semba, K. and Inoue, J. (2009) Independent stabilizations of polysomal Drg1/Dfrp1 complex and non-polysomal Drg2/Dfrp2 complex in mammalian cells. *Biochem Biophys Res Commun*, **390**, 552-556.
18. Fleischer, T.C., Weaver, C.M., McAfee, K.J., Jennings, J.L. and Link, A.J. (2006) Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev*, **20**, 1294-1307.
19. Wout, P.K., Sattlegger, E., Sullivan, S.M. and Maddock, J.R. (2009) Saccharomyces cerevisiae Rbg1 protein and its binding partner Gir2 interact on Polyribosomes with Gcn1. *Eukaryot Cell*, **8**, 1061-1071.
20. Mahajan, M.A., Park, S.T. and Sun, X.H. (1996) Association of a novel GTP binding protein, DRG, with TAL oncogenic proteins. *Oncogene*, **12**, 2343-2350.
21. Zhao, X.F. and Aplan, P.D. (1998) SCL binds the human homologue of DRG in vivo. *Biochim Biophys Acta*, **1448**, 109-114.
22. Schenker, T., Lach, C., Kessler, B., Calderara, S. and Trueb, B. (1994) A novel GTP-binding protein which is selectively repressed in SV40 transformed fibroblasts. *J Biol Chem*, **269**, 25447-25453.
23. Jakobs, A., Himstedt, F., Funk, M., Korn, B., Gaestel, M. and Niedenthal, R. (2007) Ubc9 fusion-directed SUMOylation identifies constitutive and inducible SUMOylation. *Nucleic Acids Res*, **35**, e109.
24. Eswaran, J., Bernad, A., Ligos, J.M., Guinea, B., Debreczeni, J.E., Sobott, F., Parker, S.A., Najmanovich, R., Turk, B.E. and Knapp, S. (2008) Structure of the human protein kinase MPSK1 reveals an atypical activation loop architecture. *Structure*, **16**, 115-124.
25. Chen, X. and Kumamoto, C.A. (2006) A conserved G protein (Drg1p) plays a role in regulation of invasive filamentation in Candida albicans. *Microbiology*, **152**, 3691-3700.
26. Decourty, L., Saveanu, C., Zemam, K., Hantraye, F., Frachon, E., Rousselle, J.C., Fromont-Racine, M. and Jacquier, A. (2008) Linking functionally related genes by sensitive and quantitative characterization of genetic interaction profiles. *Proc Natl Acad Sci U S A*, **105**, 5821-5826.
27. Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, **41**, 207-234.
28. Doublie, S. (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol Biol*, **363**, 91-108.
29. Sreenath, H.K., Bingman, C.A., Buchan, B.W., Seder, K.D., Burns, B.T., Geetha, H.V., Jeon, W.B., Vojtik, F.C., Aceti, D.J., Frederick, R.O. *et al.* (2005) Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. *Protein Expr Purif*, **40**, 256-267.
30. Brooks, M.A., Dziembowski, A., Quevillon-Cheruel, S., Henriot, V., Faux, C., van Tilbeurgh, H. and Seraphin, B. (2009) Structure of the yeast Pml1 splicing factor and its integration into the RES complex. *Nucleic Acids Res*, **37**, 129-143.
31. McCarthy, A.A., Brockhauser, S., Nurizzo, D., Theveneau, P., Mairs, T., Spruce,

D., Guijarro, M., Lesourd, M., Ravelli, R.B. and McSweeney, S. (2009) A decade of user operation on the macromolecular crystallography MAD beamline ID14-4 at the ESRF. *J Synchrotron Radiat*, **16**, 803-812.

32. CCP4, C.C.P. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **50**, 760-763.
33. Fortelle, E.d.L. and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods in Enzymology*, **276**, 472-494.
34. Langer, G., Cohen, S.X., Lamzin, V.S. and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc*, **3**, 1171-1179.
35. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-2132.
36. Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-255.
37. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*, **98**, 10037-10041.
38. Schrodinger, LLC. (2010).
39. Baykov, A.A., Evtushenko, O.A. and Avaeva, S.M. (1988) A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. *Anal Biochem*, **171**, 266-270.
40. Baudin-Baillieu, A., Guillemet, E., Cullin, C. and Lacroute, F. (1997) Construction of a yeast strain deleted for the TRP1 promoter and coding region that enhances the efficiency of the polymerase chain reaction-disruption method. *Yeast*, **13**, 353-356.
41. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Séraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol*, **17**, 1030-1032.
42. Wach, A., Brachat, A., Pohlmann, R. and Philippsen, P. (1994) New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast*, **10**, 1793-1808.
43. Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. *J Bacteriol*, **153**, 163-168.
44. Kushnirov, V.V. (2000) Rapid and reliable protein extraction from yeast. *Yeast*, **16**, 857-860.
45. Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M. and Séraphin, B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, **24**, 218-229.
46. Ramakrishnan, V. and White, S.W. (1992) The structure of ribosomal protein S5 reveals sites of interaction with 16S rRNA. *Nature*, **358**, 768-771.
47. Alves, V.S. and Castilho, B.A. (2005) Gir2 is an intrinsically unstructured protein that is present in Saccharomyces cerevisiae as a group of heterogeneously electrophoretic migrating forms. *Biochem Biophys Res Commun*, **332**, 450-455.

48. Sommer, K.A., Petersen, G. and Bautz, E.K. (1994) The gene upstream of DmRP128 codes for a novel GTP-binding protein of Drosophila melanogaster. *Mol Gen Genet*, **242**, 391-398.
49. Lin, B., Covalle, K.L. and Maddock, J.R. (1999) The Caulobacter crescentus CgtA protein displays unusual guanine nucleotide binding and exchange properties. *J Bacteriol*, **181**, 5825-5832.
50. Welsh, K.M., Trach, K.A., Folger, C. and Hoch, J.A. (1994) Biochemical characterization of the essential GTP-binding protein Obg of Bacillus subtilis. *J Bacteriol*, **176**, 7161-7168.
51. Ishikawa, K., Azuma, S., Ikawa, S., Morishita, Y., Gohda, J., Akiyama, T., Semba, K. and Inoue, J. (2003) Cloning and characterization of Xenopus laevis drg2, a member of the developmentally regulated GTP-binding protein subfamily. *Gene*, **322**, 105-112.
52. Yusupova, G.Z., Yusupov, M.M., Cate, J.H. and Noller, H.F. (2001) The path of messenger RNA through the ribosome. *Cell*, **106**, 233-241.
53. Davies, C., Bussiere, D.E., Golden, B.L., Porter, S.J., Ramakrishnan, V. and White, S.W. (1998) Ribosomal proteins S5 and L6: high-resolution crystal structures and roles in protein synthesis and antibiotic resistance. *J Mol Biol*, **279**, 873-888.
54. Campos-Olivas, R., Sanchez, R., Torres, D. and Blanco, F.J. (2007) Backbone assignment of the 98 kDa homotrimeric yeast PCNA ring. *J Biomol NMR*, **38**, 167.

Figures Legends:

Figure 1: **Structure of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex with sequence information**

A. A surface representation of the Tma46 C-terminal fragment (pink) enveloping Rbg1 (pale blue) is shown on the left. The individual components are shown colour-coded on the right: Rbg1 with the different domains, G-domain (pale blue), HTH and S5D2L domain (purple), TGS domain (blue) and the Tma46 C-terminal fragment (pink). The GTP binding pocket is also represented with the five G motifs coloured as orange.

B. The component sequences and secondary structure elements of the crystallized complex are represented with the G-motifs (G1-G5) given in bold letters. Domain boundaries are indicated in the same colour scheme as in figure 1.A.

Figure 2: **Electrostatic surface representation**

A. The solvent-accessible surface electrostatic potential of the Rbg1-$Tma46_{205-345}$ complex as calculated by APBS (Pymol) is shown as a surface alongside the cartoon representation. The potential is given with the negative (red) and positive (blue) contour levels in the range from -8.0- +8.0 kBT respectively. The left figure shows the positively charged surface formed partly by the G-domain, TGS and HTH-S5D2L domains.

Figure 3: **Analysis of Tma46 mutants**

A. Complementation assay for Tma46 function. The ability of plasmid-encoded Tma46 mutants to complement the growth phenotype of a triple Δ*tma46*Δ*gir2*Δ*slh1* strain was assayed by spotting serial dilution on selective plates and incubating at 30 or 37°C for 3 days. The structure of the various mutants is shown schematically on the left. WT strain indicates the original wild type parental strain without mutation. B. Mutant protein accumulation. The level of accumulation of the mutant proteins in cells shown on panel A grown at 30°C was assessed by detecting the HA tag by western blotting. Uniform loading is supported by analysis of the levels of the endogenous Stm1 protein. C. Effect of C-terminal Tma46 truncation on its binding to Rbg1 in yeast. Extracts prepared from Δ*tma46* strains carrying TAP-tagged Rbg1 and the various HA-Tma46 mutants grown at 30°C were used for immunoprecipitation on IgG beads. As control for the specificity of

the co-precipitation a wild type strain expressing wild type Tma46 tagged with HA was used. Proteins present in extracts (Input) and (co-)precipitated factors (Eluate) were analysed by western blotting. D. Effect of deletion of helices α1 and α1+α2 of Tma46 on binding to Rbg1. Samples were prepared as in panel C.

Figure 4: **GTP binding and hydrolysis**

A. Presence of GTP, GTPγS or GDP (in 4 fold excess over the protein concentration) causes an increase in melting temperature of the Rbg1-Tma46$_{205-345}$ complex in the thermal shift assay indicative of nucleotide binding. In contrast, addition of a 10 fold excess of GTPγS does not increase the melting temperature of Rbg1 alone. Note also that Rbg1 melts at a much lower temperature than Rbg1-Tma46$_{205-345}$.

B. The GTP hydrolytic activity of Rbg1/Drg1, alone and in complex with Tma46$_{205-345}$/Lerepo4$_{220-396}$ respectively, is represented as a graph with increasing substrate concentration in the x axis.

Figure 5: **Analysis of Rbg1 domains**

A. Complementation assay for Rbg1 function. The ability of plasmid-encoded Rbg1 mutants to complement the growth phenotype of a triple Δrbg1Δrbg2Δslh1 strain was assayed by spotting serial dilution on selective plates and incubating at 30 or 37°C for 4 days. The structure of the various mutants is shown schematically on the left. WT strain indicates the original wild type parental strain without mutation. B. Mutant protein accumulation. The level of accumulation of the mutant proteins in cells shown on panel A grown at 30°C was assessed by detecting the HA tag by western blotting. Uniform loading is supported by analysis of the levels of the endogenous Stm1 protein. C. Interaction between recombinant Tma46 and Rbg1 mutants. Plasmids harbouring operons encoding His6-tagged Rbg1 (wild type, point mutant or mutant deleted for specific domains) together with Tma46 (either amino-acids 154-345 or 205-345) were used to express protein in *E. coli*. Recombinant proteins purified on NiNTA agarose were detected by Coomassie staining. Organization of the different operons is shown on the left. D. Effect of C-terminal Tma46 truncation on binding to Rbg1 in yeast. Extracts prepared from Δ*rbg1* strains carrying TAP-tagged Tma46 and various HA-Rbg1 mutants

grown at 30°C were used for immunoprecipitation on IgG beads. As control for the specificity of the co-precipitation a wild type strain expressing wild type Rbg1 tagged with HA was used. Proteins present in extracts (Input) and (co-)precipitated factors (Eluate) were analysed by western blotting.

Figure 6: **RNA homopolymer binding**

A. Rbg1 and Rbg1-Tma46 complex binds to RNA: Rbg1-Tma46$_{205\text{-}345}$ complex pulled down with poly(U) beads (lane Beads) shows that the complex protein binds to poly(U). Absence of binding of Rbg1-Tma46$_{205\text{-}345}$ to RNase treated beads, demonstrate that interaction occurs through RNA. BSA, included as a negative control, demonstrates the specificity of the assay. Rbg1 and Drg1 full length proteins on their own also bound to poly(U) beads.

B. Binding can be saturated and is stable: Increasing concentrations (0.1, 1 mg/ml) of free poly(U) added to Rbg1-Tma46$_{205\text{-}345}$ before binding to the poly(U) agarose beads competed for binding demonstrating that binding occur through defined surface that can be saturated. Addition of increasing concentrations of heparin (0.5, 1, 2 and 4 mg/ml) to the washes after Rbg1-Tma46$_{205\text{-}345}$ was bound to the beads did not completely displace the protein from the beads indicating that binding is stable.

C. Contribution of individual domains of Rbg1: Rbg1$_{\Delta TGS}$ and Rbg1$_{TGS}$, both in complex with Tma46$_{DFRP}$ bound to poly(U) beads suggesting that TGS domain can bind to poly(U) but is not essential for Rbg1 RNA binding. Rbg1$_{\Delta S5D2L}$ bound to poly(U) beads while Rbg1$_{S5D2L}$ domain alone did not show any binding.

Figure 7: **Polysome association of Rbg1 mutants**

Polysomes extracts were prepared from cells expressing TAP-tagged Tma46 and HA-tagged Rbg1 (wild type or domain deletion mutants). Polysomes were resolved by density sedimentation in 10–50% sucrose gradient. The UV absorbance trace (254 nm) obtained by continuous monitoring during fractionation is shown with the position of the 40S, 60S, 80S and polysomes peaks indicated. Fractions were analysed by western blotting to detect the TAP and HA tags. A. Distribution of Tma46-TAP and wild type HA-Rbg1. B.

Distribution of Tma46-TAP and HA-Rbg1 ΔTGS. C. Distribution of Tma46-TAP and HA-Rbg1 ΔS5D2L. D. Distribution of Tma46-TAP and HA-Rbg1 ΔHTH.

Table 1: Data collection and refinement statistics of the $Rbg1_{fl}$-$Tma46_{205-345}$ complex structure

Data Collection		
Space group	$P2_12_12$	
Unit cell parameters		
a, b, c (Å)	86.2, 224.89, 84.89	
α, β, λ (°)	90.0, 90.0, 90.0	
	SeMet (peak)	Native
Wavelength (Å)	0.9795	1.0332
Completeness (%)	99.6 (99.6)	99.9 (99.9)
Mean I/σ(I)[a]	21.6 (5.9)	15.6 (4.4)
R_{meas} (%)[b]	8.3 (42.3)	8.1 (43.4)
Refinement		
Resolution (Å)		56.22-2.67 (2.74-2.67)
No. of reflections		46457
Reflections used in R_{free}		1200
R_{factor}[c]		19.7 %
R_{free}[d]		22.2 %
Stereochemistry		
Res. in favoured regions (%)		89.7
Res. in allowed regions (%)		10.2
Number of atoms		
Protein		7030
Water		77
Mean B-factor-Overall		76.935
Rmsd[e]		
Bond lengths (Å)		0.007
Bond angles (°)		1.233
Residues modeled		
Rbg1 A		2-45, 53-125, 131-369
Rbg1 B		2-91, 98-125, 133-369
Tma46 C		214-282, 302-338
Tma46 D		214-282, 320-336
Residues with missing side chain		
Rbg1 A		Lys329
Rbg1 B		Ala46, Ser47, Ser48, Ser50, Lys369
Tma46 C		Glu307
Tma46 D		Leu214, Glu215, Asp320

Numbers in parenthesis indicate the highest resolution shell statistics.

[a] Mean $[I/\sigma(I)]$ is the average of the relation between the intensity of the diffraction and the background.

[b] $R_{meas} = **\Sigma_{hkl}\ [N/(N-1)]^{1/2}\ \Sigma_i\ |I_i(hkl) - <I(hkl)>|*** / \Sigma_{hkl}\ \Sigma_i\ I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $<I(hkl)>$ are the average intensities and N is the multiplicity of reflection *hkl*.

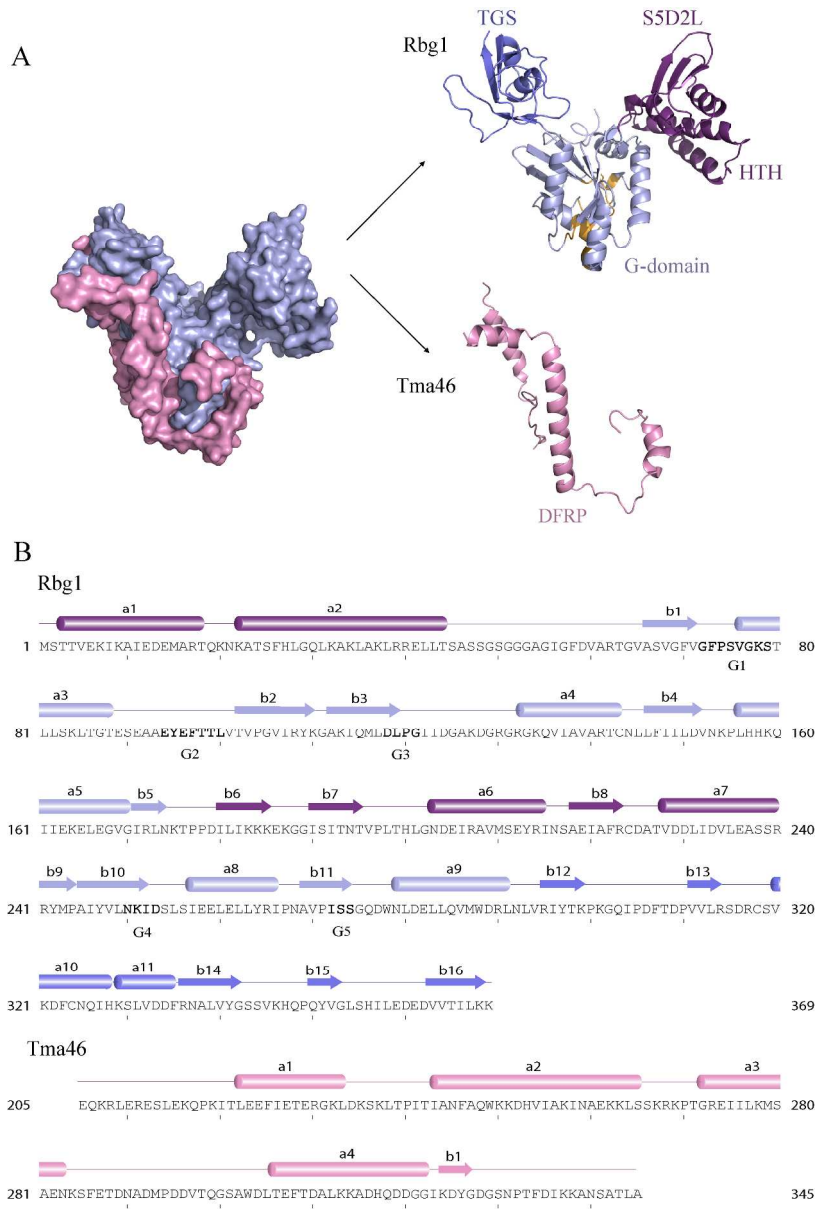
[c] $R_{factor} = \Sigma_{hkl}\ **[F_{obs}(hkl)] - [F_{calc}(hkl)]*** / \Sigma_{hkl}\ [F_{obs}(hkl)]$, where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.

[d] R_{free} corresponds to R_{factor} calculated using 2.5 % of the total reflections selected randomly and excluded during refinement.

[e] *Rmsd* is the root mean square deviation.

Table 2: Kinetic parameters for GTP hydrolysis of the Rbg1-Tma46 complex and the mutant proteins

	V_{max} *(nmol min^{-1} mg^{-1})*	K_m *(μM)*	k_{cat} *(min^{-1})*
$Rbg1_{fl}$-$Tma46_{205-345}$	0.0593 ± 0.0018	304.1 ± 32.5	0.0034
$Drg1_{fl}$ + $Lerepo4_{220-426}$	0.1154 ± 0.0037	319.5 ± 35.0	0.0074
$Rbg1_{1-294}$-$Tma46_{205-345}$ (ΔTGS)	0.0166 ± 0.0022	426.3 ± 165	0.0008
$Rbg1_{\Delta175-243+Gly}$-$Tma46_{205-345}$ (ΔS5D2L)	0.0157 ± 0.0028	221.5 ± 159.2	0.0006
$Rbg1_{fl}$-$Tma46_{239-345}$ (Tma46Δα1)	0.0370 ± 0.0012	191.2 ± 26.6	0.0020
$Rbg1_{fl}$-$Tma46_{205-345\ I241A,\ F246A}$	0.0288 ± 0.0021	428.9 ± 94.7	0.0017
$Rbg1_{fl}$-$Tma46_{205-345\ I241A}$	0.0406 ± 0.0018	186.0 ± 34.5	0.0023
$Rbg1_{fl}$-$Tma46_{205-345\ W249A\ K250E}$	0.0345 ± 0.0014	425.8 ± 52.6	0.0020

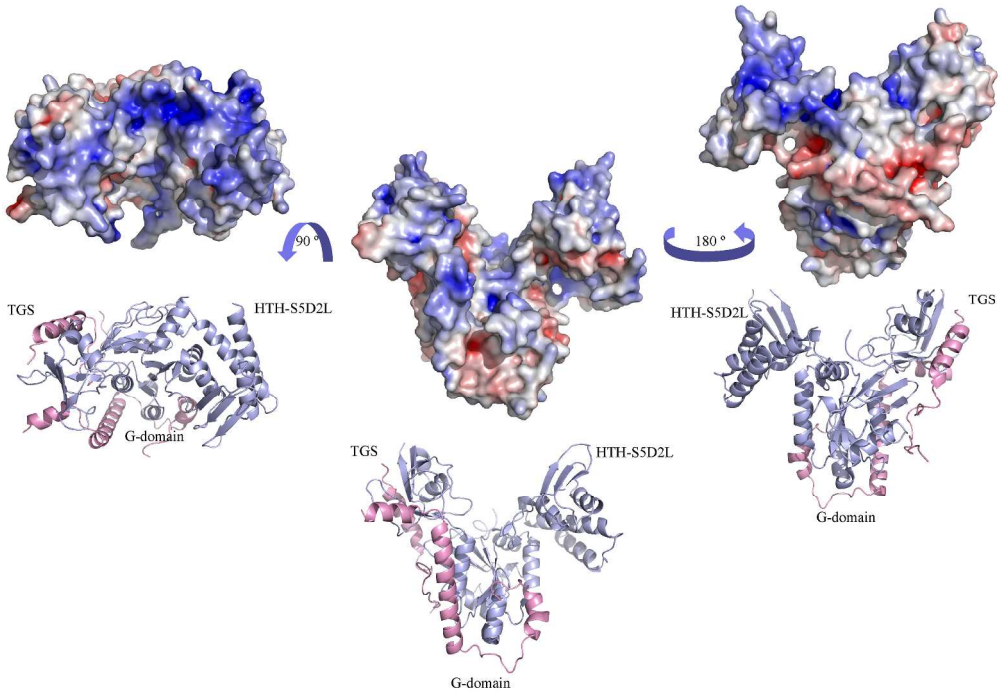


Structure of the Rbg1fl-Tma46205-345 complex with sequence information

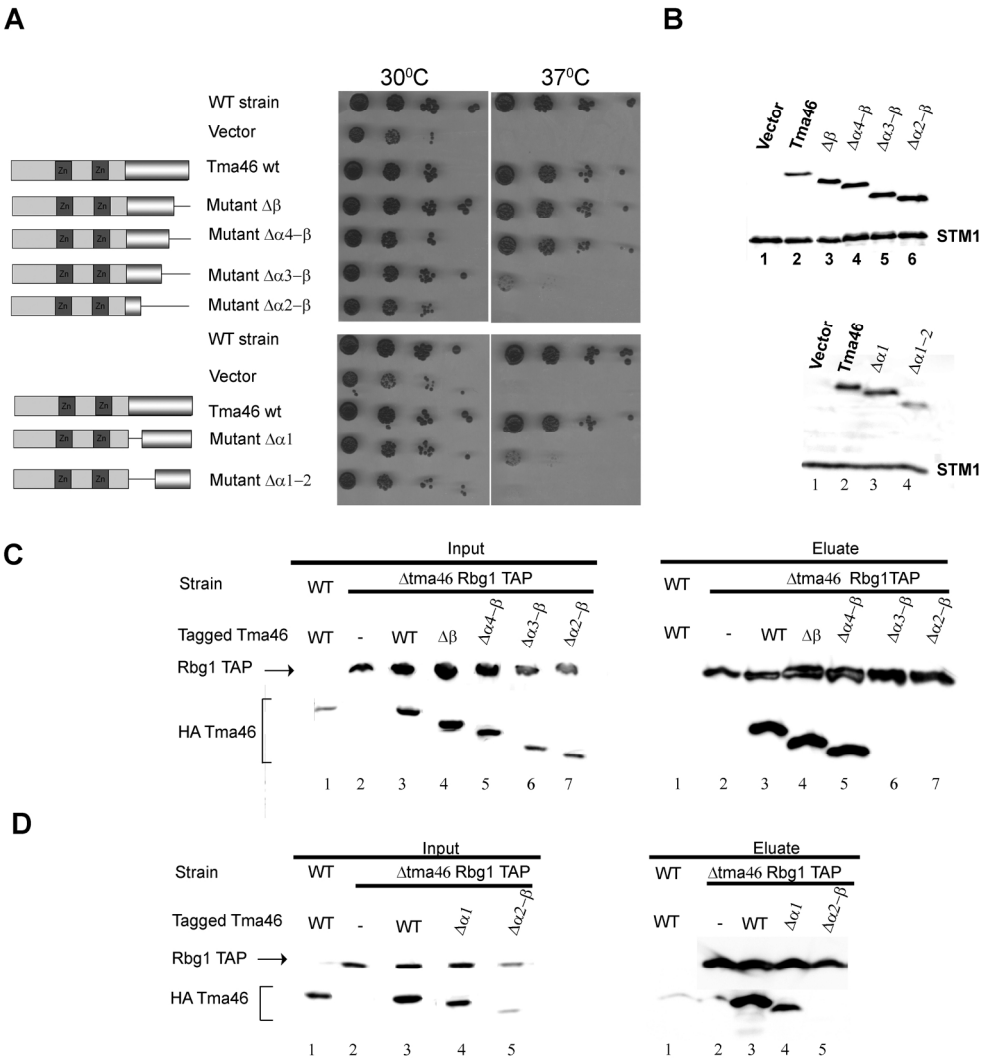
A. A surface representation of the Tma46 C-terminal fragment (pink) enveloping Rbg1 (pale blue) is shown on the left. The individual components are shown colour-coded on the right: Rbg1 with the different domains, G-domain (pale blue), HTH and S5D2L domain (purple), TGS domain (blue) and the Tma46 C-terminal fragment (pink). The GTP binding pocket is also represented with the five G motifs coloured as orange.

B. The component sequences and secondary structure elements of the crystallized complex are represented with the G-motifs (G1-G5) given in bold letters. Domain boundaries are indicated in the same colour scheme as in figure 1.A.

485x721mm (150 x 150 DPI)

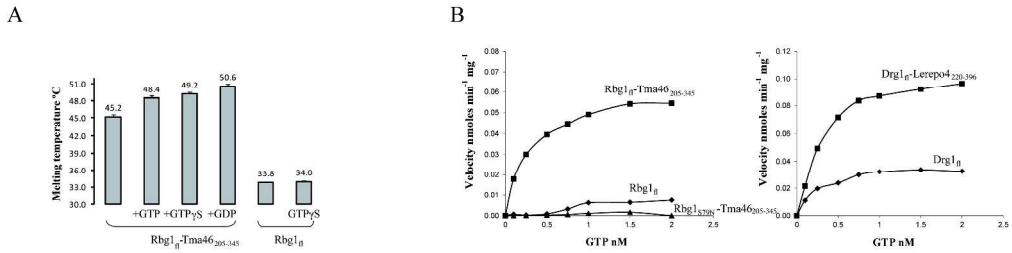


A. The solvent-accessible surface electrostatic potential of the Rbg1-Tma46205-345 complex as calculated by APBS (Pymol) is shown as a surface alongside the cartoon representation. The potential is given with the negative (red) and positive (blue) contour levels in the range from -8.0- +8.0 kBT respectively. The left figure shows the positively charged surface formed partly by the G-domain, TGS and HTH-S5D2L domains.
576x396mm (150 x 150 DPI)

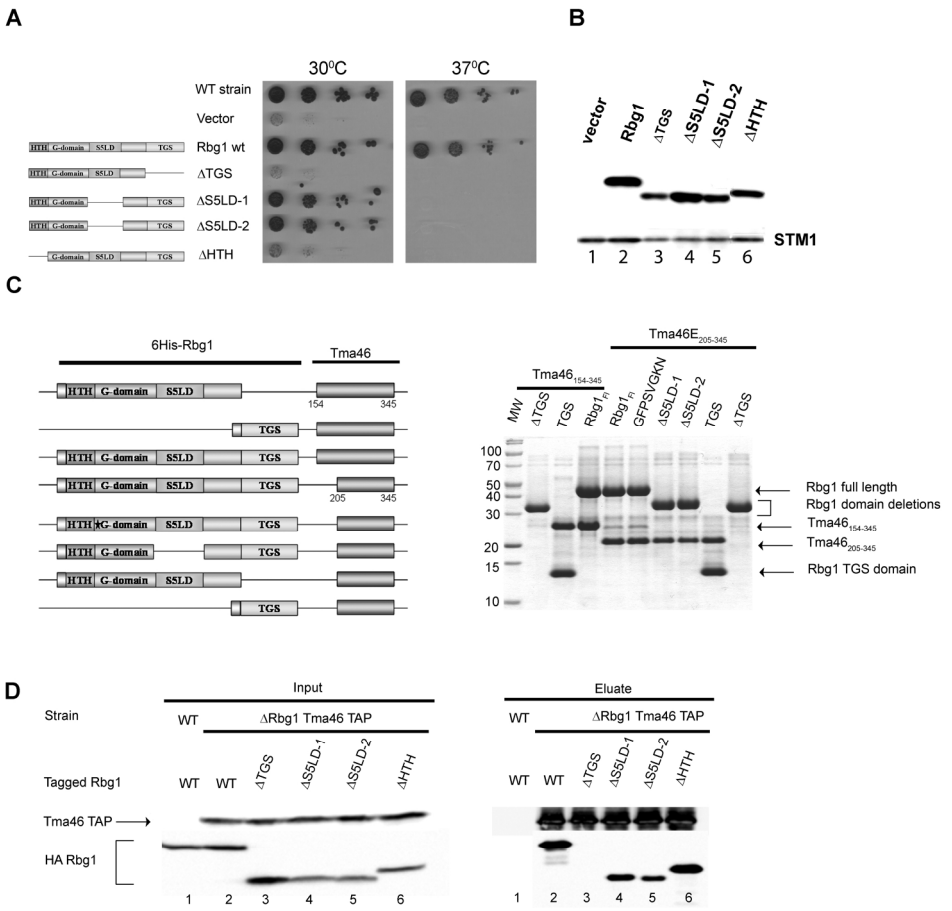


A. Complementation assay for Tma46 function. The ability of plasmid-encoded Tma46 mutants to complement the growth phenotype of a triple Δtma46Δgir2Δslh1 strain was assayed by spotting serial dilution on selective plates and incubating at 30 or 37°C for 3 days. The structure of the various mutants is shown schematically on the left. WT strain indicates the original wild type parental strain without mutation. B. Mutant protein accumulation. The level of accumulation of the mutant proteins in cells shown on panel A grown at 30°C was assessed by detecting the HA tag by western blotting. Uniform loading is supported by analysis of the levels of the endogenous Stm1 protein. C. Effect of C-terminal Tma46 truncation on its binding to Rbg1 in yeast. Extracts prepared from Δtma46 strains carrying TAP-tagged Rbg1 and the various HA-Tma46 mutants grown at 30°C were used for immunoprecipitation on IgG beads. As control for the specificity of the co-precipitation a wild type strain expressing wild type Tma46 tagged with HA was used. Proteins present in extracts (Input) and (co-)precipitated factors (Eluate) were analysed by western blotting. D. Effect of deletion of helices α1 and α1+α2 of Tma46 on binding to Rbg1. Samples were prepared as in panel C.

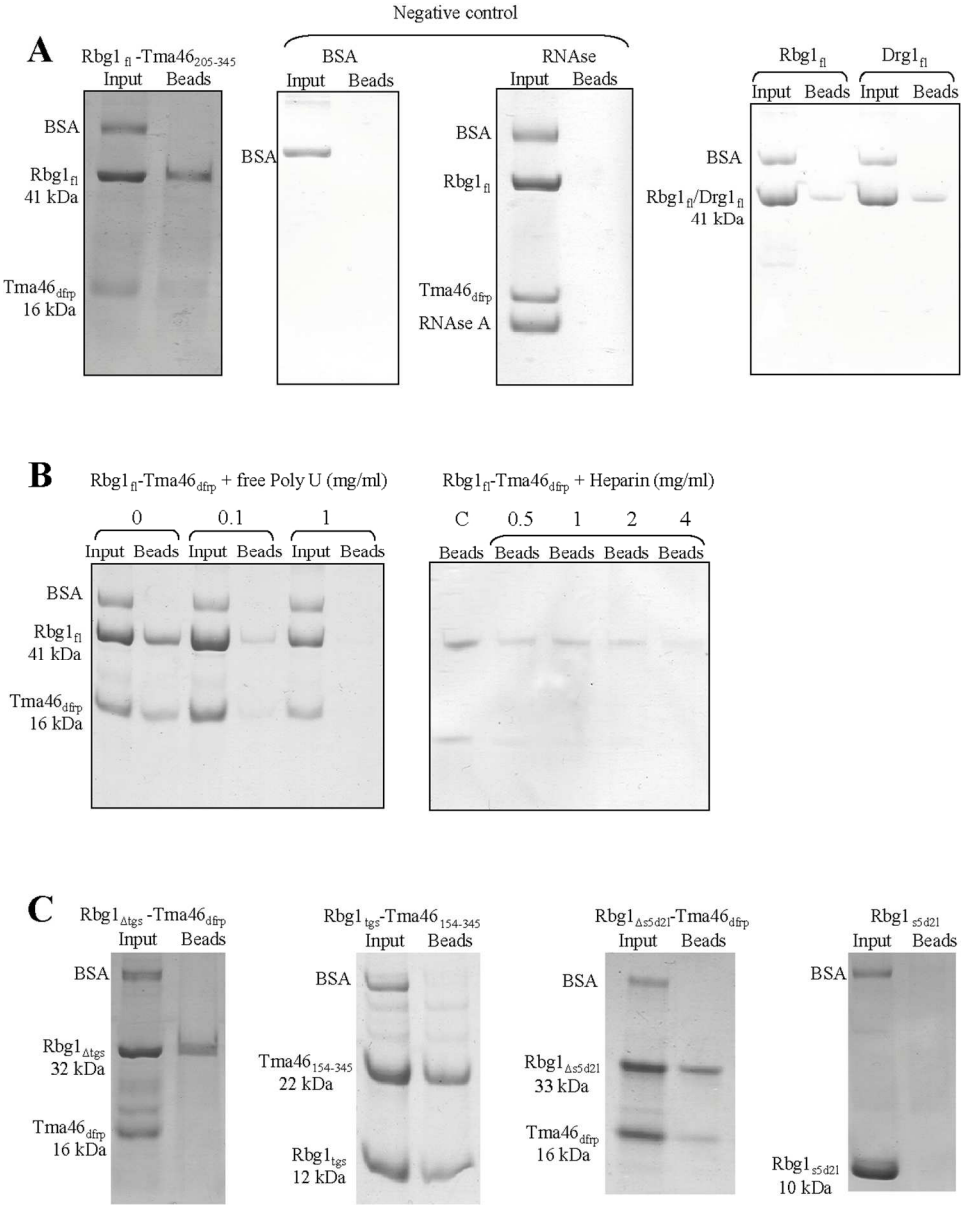
199x235mm (300 x 300 DPI)



A. Presence of GTP, GTPγS or GDP (in 4 fold excess over the protein concentration) causes an increase in melting temperature of the Rbg1-Tma46205-345 complex in the thermal shift assay indicative of nucleotide binding. In contrast, addition of a 10 fold excess of GTPγS does not increase the melting temperature of Rbg1 alone. Note also that Rbg1 melts at a much lower temperature than Rbg1-Tma46205-345.
B. The GTP hydrolytic activity of Rbg1/Drg1, alone and in complex with Tma46205-345/Lerepo4220-396 respectively, is represented as a graph with increasing substrate concentration in the x axis.
377x95mm (300 x 300 DPI)



A. Complementation assay for Rbg1 function. The ability of plasmid-encoded Rbg1 mutants to complement the growth phenotype of a triple Δrbg1Δrbg2Δslh1 strain was assayed by spotting serial dilution on selective plates and incubating at 30 or 37°C for 4 days. The structure of the various mutants is shown schematically on the left. WT strain indicates the original wild type parental strain without mutation. B. Mutant protein accumulation. The level of accumulation of the mutant proteins in cells shown on panel A grown at 30°C was assessed by detecting the HA tag by western blotting. Uniform loading is supported by analysis of the levels of the endogenous Stm1 protein. C. Interaction between recombinant Tma46 and Rbg1 mutants. Plasmids harbouring operons encoding His6-tagged Rbg1 (wild type, point mutant or mutant deleted for specific domains) together with Tma46 (either amino-acids 154-345 or 205-345) were used to express protein in E. coli. Recombinant proteins purified on NiNTA agarose were detected by Coomassie staining. Organization of the different operons is shown on the left. D. Effect of C-terminal Tma46 truncation on binding to Rbg1 in yeast. Extracts prepared from Δrbg1 strains carrying TAP-tagged Tma46 and various HA-Rbg1 mutants grown at 30°C were used for immunoprecipitation on IgG beads. As control for the specificity of the co-precipitation a wild type strain expressing wild type Rbg1 tagged with HA was used. Proteins present in extracts (Input) and (co-)precipitated factors (Eluate) were analysed by western blotting.
204x194mm (300 x 300 DPI)

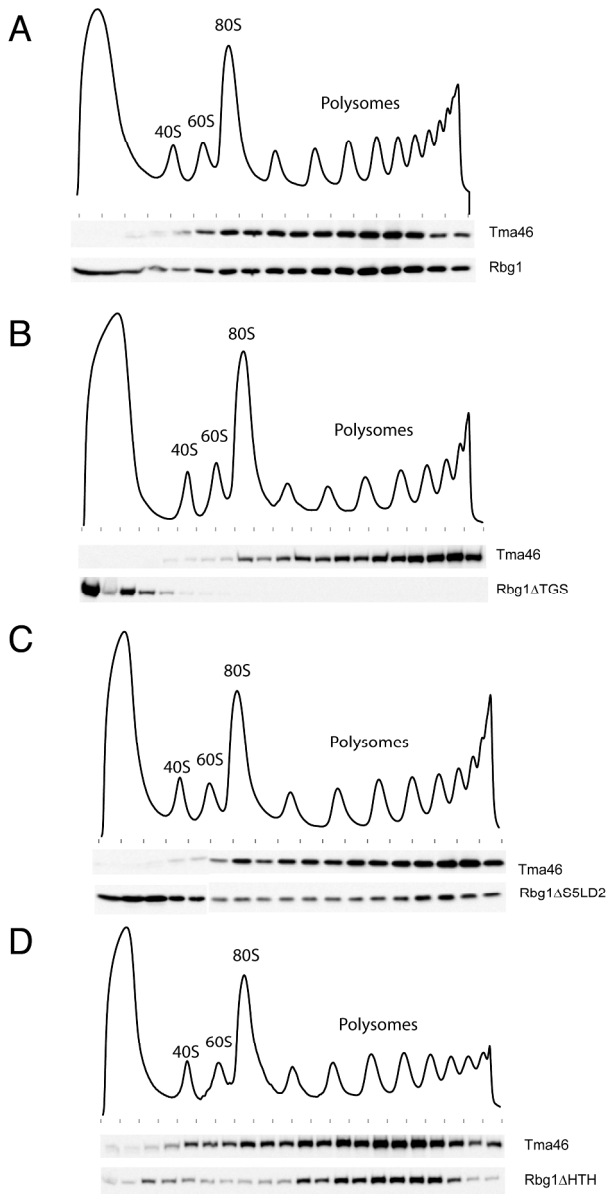


A. Rbg1 and Rbg1-Tma46 complex binds to RNA: Rbg1-Tma46205-345 complex pulled down with poly(U) beads (lane Beads) shows that the complex protein binds to poly(U). Absence of binding of Rbg1-Tma46205-345 to RNase treated beads, demonstrate that interaction occurs through RNA. BSA, included as a negative control, demonstrates the specificity of the assay. Rbg1 and Drg1 full length proteins on their own also bound to poly(U) beads.

B. Binding can be saturated and is stable: Increasing concentrations (0.1, 1 mg/ml) of free poly(U) added to Rbg1-Tma46205-345 before binding to the poly(U) agarose beads competed for binding demonstrating that binding occur through defined surface that can be saturated. Addition of increasing concentrations of heparin (0.5, 1, 2 and 4 mg/ml) to the washes after Rbg1-Tma46205-345 was bound to the beads did not completely displace the protein from the beads indicating that binding is stable.

C. Contribution of individual domains of Rbg1: Rbg1ΔTGS and Rbg1TGS, both in complex with Tma46DFRP bound to poly(U) beads suggesting that TGS domain can bind to poly(U) but is not essential for Rbg1 RNA binding. Rbg1ΔS5D2L bound to poly(U) beads while Rbg1S5D2L domain alone did not show any binding.

218x273mm (150 x 150 DPI)



Polysomes extracts were prepared from cells expressing TAP-tagged Tma46 and HA-tagged Rbg1 (wild type or domain deletion mutants). Polysomes were resolved by density sedimentation in 10–50% sucrose gradient. The UV absorbance trace (254 nm) obtained by continuous monitoring during fractionation is shown with the position of the 40S, 60S, 80S and polysomes peaks indicated. Fractions were analysed by western blotting to detect the TAP and HA tags. A. Distribution of Tma46-TAP and wild type HA-Rbg1. B. Distribution of Tma46-TAP and HA-Rbg1 ΔTGS. C. Distribution of Tma46-TAP and HA-Rbg1 ΔS5D2L. D. Distribution of Tma46-TAP and HA-Rbg1 ΔHTH.
158x283mm (300 x 300 DPI)